

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:06:48 ; Search time 40.99 Seconds
(without alignments)
46.459 Million cell updates/sec

Title: US-09-701-623c-5
Perfect score: 140
Sequence: 1 GGYQSRVTHPHLPALMRSTTKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 4959

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------------|
| 1 | 33 | 23.6 | 19 | 2 | B53145 high conductance c |
| 2 | 32 | 22.9 | 23 | 2 | F61491 seed protein ws-6 |
| 3 | 31 | 22.1 | 20 | 2 | A61093 glue protein - Cal |
| 4 | 28 | 20.0 | 24 | 2 | B44379 omega-conotoxin SV |
| 5 | 28 | 20.0 | 24 | 2 | S29749 serum albumin - do |
| 6 | 27 | 19.3 | 15 | 2 | PA0054 protein QF200017 - |
| 7 | 27 | 19.3 | 23 | 2 | S72535 probable acr-2 reg |
| 8 | 26 | 18.6 | 13 | 2 | S32551 glutathione transi |
| 9 | 26 | 18.6 | 14 | 2 | S62374 alpha-1-antichymot |
| 10 | 26 | 18.6 | 25 | 2 | A60621 somatotropin - Atl |
| 11 | 26 | 18.6 | 25 | 2 | A58647 gene GLUT4 protein |
| 12 | 25.5 | 18.2 | 25 | 2 | A58647 alpha-conotoxin p |
| 13 | 25 | 17.9 | 10 | 2 | H37196 bradykinin-potent |
| 14 | 25 | 17.9 | 21 | 2 | A56901 nerve growth facto |
| 15 | 25 | 17.9 | 22 | 2 | I37144 aspartylglycosamin |
| 16 | 25 | 17.9 | 23 | 2 | A60423 monophenol monooxy |
| 17 | 24 | 17.1 | 10 | 2 | B37196 bradykinin-potent |
| 18 | 24 | 17.1 | 12 | 2 | C30503 Ig gamma-2b chain |
| 19 | 24 | 17.1 | 13 | 2 | S28425 20K protein - rape |
| 20 | 24 | 17.1 | 19 | 2 | B61409 genome polypeptid |
| 21 | 24 | 17.1 | 20 | 2 | T44453 acetyl-CoA synthet |
| 22 | 24 | 17.1 | 20 | 2 | T01691 hypothetical prote |
| 23 | 24 | 17.1 | 21 | 2 | A60975 HMGI/2 homolog - a |
| 24 | 24 | 17.1 | 23 | 2 | S41390 p7 protein - human |
| 25 | 24 | 17.1 | 25 | 2 | A60412 xenopsin-related p |
| 26 | 24 | 17.1 | 25 | 2 | S03456 T-cell receptor al |
| 27 | 23 | 16.4 | 11 | 2 | PH0938 T-cell receptor be |
| 28 | 23 | 16.4 | 13 | 2 | S23372 alpha-conotoxin MI |
| 29 | 23 | 16.4 | 14 | 1 | NTKRM |

| | | | | | |
|----|------|------|----|---|----------------------------|
| 30 | 23 | 16.4 | 15 | 2 | PA0009 seed storage prote |
| 31 | 23 | 16.4 | 18 | 2 | S46241 NAD(P)-H-flavin oxi |
| 32 | 23 | 16.4 | 20 | 2 | proteinase inhibit |
| 33 | 23 | 16.4 | 20 | 2 | S58382 hypothetical prote |
| 34 | 23 | 16.4 | 20 | 2 | A56894 intracytalline c |
| 35 | 23 | 16.4 | 21 | 2 | PD0015 actin-binding prot |
| 36 | 23 | 16.4 | 23 | 2 | A56357 tyrosine kinase su |
| 37 | 23 | 16.4 | 24 | 2 | JC1355 brevian-1 - Frog |
| 38 | 23 | 16.4 | 24 | 2 | PC2305 gaegurin 6 - Kore |
| 39 | 23 | 16.4 | 24 | 2 | E45087 cysteine proteinas |
| 40 | 23 | 16.4 | 25 | 2 | A60704 serine proteinase |
| 41 | 22.5 | 16.1 | 23 | 2 | B61079 listeriolysin O - |
| 42 | 22 | 15.7 | 13 | 2 | JH0460 corticostatic pept |
| 43 | 22 | 15.7 | 13 | 2 | A54326 glandular kallikre |
| 44 | 22 | 15.7 | 15 | 2 | B41868 hypothetical prote |
| 45 | 22 | 15.7 | 16 | 2 | S30384 |

ALIGNMENTS

RESULT 1
B53145 high conductance calcium-activated potassium channel, maxi-K channel - bovine (fragme
C:Species: Bos primigenius taurus (cattle)
C>Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C:Accession: B53145
R:Knaus, H.G.; Garcia-Calvo, M.; Kaczorowski, G.J.; Garcia, M.L.
J. Biol. Chem. 269, 3921-3924, 1994
A>Title: Subunit composition of the high conductance calcium-activated potassium chan
A:Reference number: A53145; MUID:94140798
A:Accession: B53145
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-19 <RNA>
A>Note: sequence extracted from NCBI backbone (NCBIP:144547)

Query Match 23.6%; Score 33; DB 2; Length 19;
Best Local Similarity 54.5%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 EFGQSVTHPH 13
DB 6 EFGQSVTHPH 16

RESULT 2
F61491 seed protein ws-6 - winged bean (fragment)
C:Species: Psophocarpus tetragonolobus (winged bean)
C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 03-May-1996
C:Accession: F61491
J. Hirano, H.
J. Protein Chem. 8, 115-130, 1989
A>Title: Microsequence analysis of winged bean seed proteins electrobotted from two-
A:Reference number: A61491; MUID:89351606
A:Accession: F61491
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-23 <HIR>
C:Superfamily: pathogenesis-related protein
C:Keywords: seed

Query Match 22.9%; Score 32; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 1.8e+02;
Matches 6; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 4 TYQSRVTHPHALMR 21
DB 4 TYQSRVTHPHALMR 21

RESULT 3
A61093
glue protein - California mussel (fragments)
N:Alternate names: adhesive polyphenolic protein
C:Species: Mytilus californianus (California mussel)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
C:Accession: A61093
R:Matte, J.H.
J. Comp. Physiol. B 156, 491-496, 1986
A:Title: Mussel glue from Mytilus californianus Conrad: a comparative study.
A:Reference number: A61093; MUID:86279063
A:Accession: A61093
A:Molecule type: protein
A:Residues: 1-20 <MAI>
A:Note: the amino terminal residue was tentatively identified as Ser
A:Note: 12-ser was also found
C:Comment: This glue protein is a component of the adhesive plaque of the byssus, the fi
C:Keywords: hydroxyproline; tandem repeat
F:4.14/Modified site: 3-hydroxyproline (Pro) #status experimental
F:5.9.15.19/Modified site: 4-hydroxyproline (Pro) #status experimental
F:7.17/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental
F:13/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) (partial) #status experimental

Query Match
Best Local Similarity 44.4%; Score 31; DB 2; Length 20;
Pred. No. 2.2e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 4 TYQSRVTHP 12
||: ||: ||:
DB 6 TYKPKRTYP 14

RESULT 4
B44379
omega-conotoxin SV1A - cone shell (Conus striatus)
N:Alternate names: SNX-157
C:Species: Conus striatus (striated cone)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 23-May-1997
C:Accession: B44379
R:Ramilo, C.A.; Zafarani, G.C.; Nadaadi, L.; Hammerland, L.G.; Yoshikami, D.; Gray, W.H.
Biochemistry 31, 9919-9926, 1992
A:Title: Novel alpha- and omega-conotoxins from Conus striatus venom.
A:Reference number: A44379; MUID:93003172
A:Accession: B44379
A:Molecule type: protein
A:Residues: 1-24 <RAM>
A:Cross-references: CAS:137706-81-1
A:Experimental source: Venom
A:Note: sequence extracted from NCBI backbone (NCBI:P116001); structure confirmed by che
C:Comment: This omega-conotoxin blocks presynaptic calcium channels.
C:Superfamily: omega-conotoxin
C:Keywords: acetylcholine release inhibition; amidated carboxyl end; calcium channel inh
F:1.15.8-18.14-23/Disulfide bonds: #status predicted
F:24/Modified site: amidated carboxyl end (Thr) #status experimental

Query Match
Best Local Similarity 20.0%; Score 28; DB 2; Length 24;
Pred. No. 7.5e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 CGETYSRV 10
||: ||: ||:
DB 15 CGCRTRGKCT 24

RESULT 5
S29749
serum albumin - dog
C:Species: Canis lupus familiaris (dog)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999

C:Accession: S29749
R:Dixon, J.W.; Sarkar, B.
J. Biol. Chem. 249, 5872-5877, 1974
A:Title: Isolation, amino acid sequence and copper(II)-binding properties of peptide
A:Reference number: S29749; MUID:75011422
A:Accession: S29749
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-24 <DIX>
C:Superfamily: serum albumin; serum albumin repeat homology

Query Match
Best Local Similarity 20.0%; Score 28; DB 2; Length 24;
Pred. No. 7.5e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 3 ETYQSRVTH 11
||: ||: ||:
DB 1 EAYKSEIAH 9

RESULT 6
PA0054
protein QF200017 - fungus (Fusarium sporotrichoides) (fragment)
C:Species: Fusarium sporotrichoides
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Mar-2001
C:Accession: PA0054
R:Chow, L.P.; Fukaya, N.; Sugiyura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JRPD, October 1994
A:Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporot
A:Reference number: PA0054
A:Accession: PA0054
A:Molecule type: protein
A:Residues: 1-15 <CHO>

Query Match
Best Local Similarity 19.3%; Score 27; DB 2; Length 15;
Pred. No. 6.5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 GETYSRV 9
||: ||: ||:
DB 3 GETQQTQV 10

RESULT 7
S72535
probable acr-2 regulatory leader protein - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 21-Jul-2000
C:Accession: S72535
R:Akiyama, M.; Nakashima, H.
Biochim. Biophys. Acta 1307, 187-192, 1996
A:Title: Molecular cloning of the acr-2 gene which controls acriflavine sensitivity 1
A:Reference number: S72535; MUID:96283814
A:Accession: S72535
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-23 <AKT>
A:Cross-references: EMBL:D45893; NID:91754593; PIDN:BA08306.1; PID:q1754594

Query Match
Best Local Similarity 19.3%; Score 27; DB 2; Length 23;
Pred. No. 1e+03;
Matches 4; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 12 PHLPALMRSTTKC 25
||: ||: ||:
DB 7 PYOPRMIRLEWSSC 20

RESULT 8
S32551

glutathione transferase (EC 2.5.1.18) mu (isoform pI 6.4) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 05-Jan-1996
C:Accession: S32551; S32550
R:Singhal, S.S.; Saxena, M.; Ahmad, H.; Awasthi, Y.C.
Biochim. Biophys. Acta 1116, 137-146, 1992
A:Title: Glutathione S-transferases of mouse liver: sex-related differences in the expression
A:Reference number: S32548; MUID:92256466
A:Accession: S32551
A:Molecule type: protein
A:Residues: 1-13 <SIN1>
A:Experimental source: female
A:Accession: S32550
A:Molecule type: protein
A:Residues: 1-13 <SIN2>
A:Experimental source: male
C:Keywords: transferase

Query Match 18.6%; Score 26; DB 2; Length 13;
Best Local Similarity 80.0%; Pred. No. 8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 RVTTP 12
1:1111
DB 9 RVTTP 13

RESULT 9
S62374
alpha-1-antichymotrypsin - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 31-Mar-1997
C:Accession: S62374
R:Tsuda, M.; Sei, Y.; Ohkubo, T.; Yamamura, M.; Kamiguchi, H.; Akatsuka, A.; Tsuda, T.;
Eur. J. Biochem. 235, 821-827, 1996
A:Title: The defective secretion of a naturally occurring alpha-1-antichymotrypsin variant
A:Reference number: S62374; MUID:96184564
A:Accession: S62374
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-14 <TSU>

Query Match 18.6%; Score 26; DB 2; Length 14;
Best Local Similarity 50.0%; Pred. No. 8.6e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 YOSRVTTP 12
1:1111
DB 4 FMSKVTNP 11

RESULT 10
A60621
somatotropin - Atlantic salmon (fragment)
N:Alternate names: growth hormone
C:Species: Salmo salar (Atlantic salmon)
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 30-Sep-1993
C:Accession: A60621
R:Skibeli, V.; Andersen, O.; Gautvik, K.M.;
Gen. Comp. Endocrinol. 80, 333-344, 1990
A:Title: Purification and characterization of Atlantic salmon growth hormone and evidence
A:Reference number: A60621; MUID:91146880
A:Accession: A60621
A:Molecule type: protein
A:Residues: 1-25 <SKT>
A>Note: this protein displayed charge heterogeneity from variable levels of phosphorylation
C:Superfamily: prolactin
C:Keywords: hormone; phosphoprotein; pituitary

Query Match 18.6%; Score 26; DB 2; Length 25;

Best Local Similarity 62.5%; Pred. No. 1.6e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 SRVTHPHL 14
1:1111
DB 12 NRVOHLHL 19

RESULT 11
I67422
gene GLUT4 protein - rat (fragment)
C:Species: Rattus sp. (rat)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 04-Sep-1998
C:Accession: I67422
R:Olson, A.L.; Edgington, N.P.; Moye-Rowley, W.S.; Passin, J.E.
Endocrinology 136, 1962-1968, 1995
A:Title: Characterization of 5'-heterogeneity of the rat GLUT4/muscle-adipose glucose
A:Reference number: I53293; MUID:95237122
A:Accession: I67422
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-25 <RES>
A:Cross-references: GB:S77784; MUID:9998977
C:Genetics:
A:Gene: GLUT4
C:Superfamily: glucose transport protein

Query Match 18.6%; Score 26; DB 2; Length 25;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GERTYOSRVT 10
1:1111
DB 13 GEPPQORVT 21

RESULT 12
A58647
alpha-conotoxin P1YA [validated] - cone shell (Conus purpurascens)
C:Species: Conus purpurascens (purple cone)
C:Date: 31-Oct-1997 #sequence_revision 07-Nov-1997 #text_change 15-Sep-2000
C:Accession: A58647
R:Hopkins, C.; Grille, M.; Miller, C.; Shon, K.J.; Cruz, L.J.; Gray, W.R.; Dykert, J.
J. Biol. Chem. 270, 22361-22367, 1995
A:Title: A new family of Conus peptides targeted to the nicotinic acetylcholine receptor
A:Reference number: A58647; MUID:95403432
A:Accession: A58647
A:Molecule type: protein
A:Residues: 1-25 <HOP>
R:Han, K.H.; Hwang, K.J.; Kim, S.M.; Kim, S.K.; Gray, W.R.; Olivera, B.M.; Rivier, J.
submitted to the Brookhaven Protein Data Bank, December 1996
A:Reference number: A67666; PDB:1P1P
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, res1
R:Han, K.H.; Hwang, K.J.; Kim, S.M.; Kim, S.K.; Gray, W.R.; Olivera, B.M.; Rivier, J.
Biochemistry 36, 1669-1677, 1997
A:Title: NMR structure determination of a novel conotoxin, [Pro 7,13] alpha A-conotoxin
A:Reference number: A58646; MUID:97200721
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
C:Superfamily: unassigned conotoxins
C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; hydroxyproline;
F:2-16,3-11,14-23/disulfide bonds: status experimental
F:7,13/modified site: 4-hydroxyproline (Pro) (partial) status experimental
F:20/modified site: 4-hydroxyproline (Pro) status experimental
F:25/modified site: amidated carboxyl end (Gln) status experimental

Query Match 18.2%; Score 25.5; DB 2; Length 25;
Best Local Similarity 41.7%; Pred. No. 1.9e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 1; Gaps 0;

QY 1 CGERTYOSRVTTP 12
1:1111

Db 3 CG-SYNNACHP 13

RESULT 13

H37196
bradykinin-potentiating peptide 8 - island jararaca
C:Species: Bothrops insularis (Island Jararaca)
C:Date: 14-Feb-1992 #sequence_revision 01-Dec-1992 #text_change 05-Aug-1994
C:Accession: H37196
R:Clntra, A.C.O.; Vleira, C.A.; G1911o, J.R.
J. Protein Chem. 9, 221-227, 1990
A:Title: Primary structure and biological activity of bradykinin potentiating peptides
A:Reference number: A37196; MUID:90351557
A:Accession: H37196
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 1-10 <CIN>
C:Keywords: pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match

Best Local Similarity 17.9%; Score 25; DB 2; Length 10;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 11 HPNLP 15

DB 5 HPNIP 9

RESULT 14

A56901
nerve growth factor - puff adder (fragment)
C:Species: Bitis arietans (puff adder)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 20-Jun-2000
C:Accession: A56901
R:Smith, P.J.; Brandt, W.F.; Stickells, B.J.; von Holt, C.
Comp. Biochem. Physiol. B 103, 975-980, 1992
A:Title: Bitis arietans nerve growth factor is a disulphide-linked homodimer.
A:Reference number: A56901; MUID:93121763
A:Accession: A56901
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 1-21 <SMI>
A:Note: sequence extracted from NCBI backbone (NCBIP:122566)
C:Keywords: disulfide bond; homodimer; venom

Query Match

Best Local Similarity 17.9%; Score 25; DB 2; Length 21;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 3 EYQSRVTHP 12

DB 7 EYQSRVTHP 16

RESULT 15

I37144
aspartylglycosaminuria - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C:Accession: I37144
R:Park, H.; Velteise, M.B.; Fensom, A.H.; Fisher, K.J.; Aronson, N.N.
Biochem. J. 290, 735-741, 1993
A:Title: Characterization of three alleles causing aspartylglycosaminuria: two from a Br
A:Reference number: I37144; MUID:93207523
A:Accession: I37144
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-22 <RES>
A:Cross-references: EMBL:X73071; NID:g312227; PIDN:CA51529.1; PID:g312228
C:Genetics:

A:Gene: AGU

Query Match

Best Local Similarity 17.9%; Score 25; DB 2; Length 22;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 TYQSRVTH 11

DB 5 TFSRRVSH 12

Search completed: March 4, 2002, 13:11:39
Job time: 291 sec

11/13/2001 10:10:10 AM

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:10:13 ; Search time 24.21 Seconds
(without alignments)
37.861 Million cell updates/sec

Title: US-09-701-623C-5
Perfect score: 140
Sequence: 1 CGETYQSRVTHPLRALMRSTYKC 25

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1446

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 28 | 20.0 | 24 | 1 CXOA_CONST | P28880 conus stria |
| 2 | 26 | 18.6 | 21 | 1 CFPA_TREPH | P56738 treponema p |
| 3 | 25.5 | 18.2 | 25 | 1 CXAL_CONPU | P55983 conus purpu |
| 4 | 25 | 17.9 | 10 | 1 BPP8_BOVIN | P30426 bothrops in |
| 5 | 25 | 17.9 | 20 | 1 M117_BOVIN | P35451 bos taurus |
| 6 | 24 | 17.1 | 10 | 1 BPP2_BOVIN | P30422 bothrops in |
| 7 | 24 | 17.1 | 15 | 1 MCA2_RHOOP | P56870 rhodococcus |
| 8 | 24 | 17.1 | 19 | 1 HI70_RAT | P21794 rattus norv |
| 9 | 23 | 16.4 | 14 | 1 CXAL_CONMA | P01521 conus megus |
| 10 | 23 | 16.4 | 20 | 1 UN05_PINPS | P81674 pinus pinas |
| 11 | 23 | 16.4 | 21 | 1 FA71_TETPY | P81424 tetrahymena |
| 12 | 23 | 16.4 | 23 | 1 SODM_RANCA | P36215 rana catesb |
| 13 | 23 | 16.4 | 24 | 1 BRL_RANBP | P32423 rana brevip |
| 14 | 23 | 16.4 | 24 | 1 GAE5_RANRU | P80400 rana rugosa |
| 15 | 22 | 15.7 | 10 | 1 MALE_KLEPN | Q05584 klebsiella |
| 16 | 22 | 15.7 | 16 | 1 YMOR_PSEPU | Q02210 pseudomonas |
| 17 | 22 | 15.7 | 20 | 1 MIF_PIG | P80928 sus scrofa |
| 18 | 22 | 15.7 | 21 | 1 TERT_APIME | P56587 apis mellif |
| 19 | 22 | 15.7 | 24 | 1 BRL1_RANES | P40835 rana escul |
| 20 | 21.5 | 15.4 | 22 | 1 TX12_TRIWA | P24335 trimeresuru |
| 21 | 21 | 15.0 | 13 | 1 NEUT_CAVPO | P32580 cavia porce |
| 22 | 21 | 15.0 | 14 | 1 CXAL_CONCN | P56973 conus conso |
| 23 | 21 | 15.0 | 17 | 1 MAOX_CHICK | Q92060 gallus gall |
| 24 | 21 | 15.0 | 17 | 1 APID_BOMPA | P81464 bombus pasc |
| 25 | 21 | 15.0 | 18 | 1 PH1_PERAM | P82694 periplaneta |
| 26 | 21 | 15.0 | 19 | 1 CXA2_CONST | P28879 conus stria |
| 27 | 21 | 15.0 | 19 | 1 FRHG_METBA | P80491 methanosarc |
| 28 | 21 | 15.0 | 20 | 1 CS21_STSTR | P81621 streptococc |
| 29 | 21 | 15.0 | 20 | 1 TL18_SPTOL | P82536 spinacia ol |
| 30 | 21 | 15.0 | 21 | 1 THAN_PODMA | P55788 podisus mac |
| 31 | 21 | 15.0 | 24 | 1 FEDG_AMEYE | P80707 amycolatops |
| 32 | 21 | 15.0 | 24 | 1 Y3KD_NEUCR | P22702 neurospora |
| 33 | 20 | 14.3 | 12 | 1 UR2_POU5P | P61022 polyodon sp |

| | | | | | |
|----|------|------|----|--------------|--------------------|
| 34 | 20 | 14.3 | 13 | 1 CXAL_CONST | P15471 conus stria |
| 35 | 20 | 14.3 | 15 | 1 ACEA_ACICA | P28467 acinetobact |
| 36 | 20 | 14.3 | 15 | 1 CXAL_CONGE | P01519 conus geogr |
| 37 | 20 | 14.3 | 15 | 1 SODM_ENTAE | P22799 enterobacte |
| 38 | 20 | 14.3 | 15 | 1 UN01_PINPS | P81106 pinus pinas |
| 39 | 20 | 14.3 | 16 | 1 CXAL_CONEP | P56638 conus episc |
| 40 | 20 | 14.3 | 18 | 1 A2M_OCTVU | P30800 octopus vul |
| 41 | 20 | 14.3 | 20 | 1 SODE_PASPI | P81527 pasteurella |
| 42 | 20 | 14.3 | 24 | 1 LPER_STRFR | P45440 streptomyce |
| 43 | 20 | 14.3 | 25 | 1 GBX1_MOUSE | P82976 mus musculu |
| 44 | 20 | 14.3 | 25 | 1 RL41_SCHPO | Q9Y710 schizosacch |
| 45 | 19.5 | 13.9 | 17 | 1 PROP_SALTY | P40862 salmonella |

ALIGNMENTS

RESULT 1

| ID | CXOA_CONST | STANDARD | PRT | 24 AA. |
|----|--------------------------------------------------------------------------------------------------------|----------|-----|--------|
| AC | P28880; | | | |
| DT | 01-DEC-1992 (Rel. 24, Created) | | | |
| DT | 01-DEC-1992 (Rel. 24, Last sequence update) | | | |
| DT | 01-DEC-1992 (Rel. 24, Last annotation update) | | | |
| DE | OMEGA-CONOTOXIN SVIA. | | | |
| OS | Conus striatus (Striated cone). | | | |
| OC | Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda; | | | |
| OC | Neogastropoda; Conoidea; Conidae; Conus. | | | |
| OX | NCBI_TaxID=6493; | | | |
| RN | [1] | | | |
| RP | SEQUENCE, AND SYNTHESIS. | | | |
| RC | TISSUE=Venom; | | | |
| RX | MEDLINE=93003172; PubMed=1390774; | | | |
| RA | Ramilo C., Zafaralla G.C., Nadasdi L., Hammerland L.G., Yoshikami D., | | | |
| RA | Gray W.R., Kristipati R., Ramachandran J., Miljanich G., Olivera B.M., | | | |
| RA | Cruz L.J.; | | | |
| RT | "Novel alpha- and omega-conotoxins from Conus striatus venom." | | | |
| RL | Biochemistry 31:9919-9926(1992). | | | |
| CC | -1- FUNCTION: OMEGA-CONOTOXINS ACT AT PRESYNAPTIC MEMBRANES, THEY BIND AND BLOCK THE CALCIUM CHANNELS. | | | |
| DR | PIR: B44379; B44379. | | | |
| KW | Presynaptic neurotoxin; Calcium channel inhibitor; Venom; | | | |
| KW | Amidation; Hydroxylation. | | | |
| FT | DISULFID 1 15 | | | |
| FT | DISULFID 8 18 | | | |
| FT | DISULFID 14 23 | | | |
| FT | MOD_RES 7 7 | | | |
| FT | MOD_RES 24 24 | | | |
| SQ | SEQUENCE 24 AA; 2485 MW; B29EFC982ABBE644 CRC64; | | | |

Query Match 20.0%; Score 28; DB 1; Length 24;
Best Local Similarity 40.0%; Pred. No. 3e+02;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYQSRVT 10
|| : : |

Db 15 CGRCYRGKCT 24

RESULT 2

| ID | CFPA_TREPH | STANDARD | PRT | 21 AA. |
|----|-------------------------------------------------------|----------|-----|--------|
| AC | P56738; | | | |
| DT | 30-MAY-2000 (Rel. 39, Created) | | | |
| DT | 30-MAY-2000 (Rel. 39, Last sequence update) | | | |
| DT | 30-MAY-2000 (Rel. 39, Last annotation update) | | | |
| DE | CYTOPLASMIC FILAMENT PROTEIN A (FRAGMENT). | | | |
| GN | CFPA. | | | |
| OS | Treponema phagedenis. | | | |
| OC | Bacteria; Spirochaetales; Spirochaetaceae; Treponema. | | | |
| OX | NCBI_TaxID=162; | | | |
| RN | [1] | | | |

RP SEQUENCE.
RC STRAIN=KAZAN 5;
RX MEDLINE=96236033; PubMed=8655496;
RA You Y., Elmore S., Colton L.L., Mackenzie C., Stoops J.K.,
RA Weinstock G.M., Norris S.J.;
RT "Characterization of the cytoplasmic filament protein gene (cfpa) of
RT Treponema pallidum subsp. pallidum";
RL J. Bacteriol. 178:3177-3187(1996).
CC -1- FUNCTION: COMPONENT OF THE CYTOPLASMIC FILAMENTS THAT RUN THE
CC LENGTH OF THE ORGANISM JUST UNDERNEATH THE CYTOPLASMIC MEMBRANE.
CC -1- SUBCELLULAR LOCATION: AN ARRAY OF 4 TO 6 FILAMENTS LIE IN CLOSE
CC APPosition TO THE INNER MEMBRANE AND ARE ALWAYS LOCALIZED DIRECTLY
CC UNDERNEATH THE CORRESPONDING GROUP OF PERIPLASMIC FLAGELLA.
KW Structural protein; Antigen.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2231 MW; 574604B4FFC2D017 CRC64;

Query Match 18.6%; Score 26; DB 1; Length 21;
Best Local Similarity 50.0%; Pred. No. 5.3e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 9 VTHPLPRAL 18
Db 11 VFHPEKPSAV 20
| | | | |
| | | | |

RESULT 3
ID CXAA_CONPU STANDARD; PRT; 25 AA.
AC P55963;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ALPHA-A CONOTOXIN PIVA.
OS Conus purpurascens (purple cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=41690;
RN [1]
RP TISSUE=Venom;
RC MEDLINE=95403432; PubMed=7673220;
RA Hopkins C., Grille M., Miller C., Shon K.-J., Cruz L.J., Gray W.R.,
RA Dykert J., Rivier J., Yoshikami D., Olivera B.M.;
RT "A new family of Conus peptides targeted to the nicotinic
RT acetylcholine receptor";
RL J. Biol. Chem. 270:22361-22367(1995).
RN [2]
RP STRUCTURE BY NMR.
RX MEDLINE=97200721; PubMed=9048550;
RA Han K.-H., Hwang K.-J., Kim S.-M., Kim S.-K., Gray W.R., Olivera B.M.,
RA Rivier J., Shon K.-J.;
RT "NMR structure determination of a novel conotoxin, [Pro 7,13] alpha
RT A-conotoxin PIVA";
RL Biochemistry 36:1569-1577(1997).
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
DR PDB; IPIP; 07-JUL-97.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Hydroxylation; Venom; 3D-structure.
FT DISULFID 2 16
FT DISULFID 3 11
FT DISULFID 14 23
FT MOD_RES 7 7 HYDROXYLATION.
FT MOD_RES 13 13 HYDROXYLATION.
FT MOD_RES 20 20 HYDROXYLATION.
FT MOD_RES 25 25 AMIDATION.
SQ SEQUENCE 25 AA; 2608 MW; 9E2147898D697640 CRC64;

Query Match 18.2%; Score 25.5; DB 1; Length 25;

Best Local Similarity 41.7%; Pred. No. 7.6e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Qy 1 CGEYQSRVTHP 12
Db 3 CG-SYPNAAACP 13
| | : | |
| | : | |

RESULT 4
ID BPP8_BOTIN STANDARD; PRT; 10 AA.
AC P30426;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE BRADYKININ-POTENTIATING PEPTIDE S5,1 (ANGIOTENSIN-CONVERTING
DE ENZYME INHIBITOR).
OS Bothrops insularis (Island jararaca) (Quelma jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=90351557; PubMed=2386615;
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
RT "Primary structure and biological activity of bradykinin potentiating
RT peptides from Bothrops insularis snake venom";
RL J. Protein Chem. 9:221-227(1990).
CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
DR PIR; H37196; H37196.
KW Hypotensive agent; Venom.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 10 AA; 1173 MW; 2FF835545761F6D8 CRC64;

Query Match 17.9%; Score 25; DB 1; Length 10;
Best Local Similarity 60.0%; Pred. No. 3.4e+02;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 11 HPHLP 15
Db 5 HPNIP 9
| | : | |
| | : | |

RESULT 5
ID M17_BOVIN STANDARD; PRT; 20 AA.
AC P35451;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE 17 KDA MILK GLYCOPROTEIN (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Milk;
RX MEDLINE=93308294; PubMed=8320368;
RA Soerensen E.S., Petersen T.E.;
RT "Purification and characterization of three proteins isolated from
RT the proteose peptone fraction of bovine milk";
RL J. Dairy Res. 60:189-197(1993).
CC -1- PTM: N-GLYCOSYLATED.
CC -1- SIMILARITY: TO CAMEL WHEY PROTEIN.
KW Glycoprotein; Milk.

FT NON_TER 1 1
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2233 MW; 4CCA589404C62C27 CRC64;

Query Match 17.9%; Score 25; DB 1; Length 20;
Best Local Similarity 26.7%; Pred. No. 7.1e+02;
Matches 4; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 6 QSRVTHPLRALMR 20
| : | | : : :
Db 4 QPSONPKPLSLK 18

RESULT 6

ID BPP2_BOTIN STANDARD; PRT; 10 AA.
AC P30422;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE BRADYKININ-POTENTIATING PEPTIDE S4,3,1 (10C) (ANGIOTENSIN-CONVERTING ENZYME INHIBITOR).
OS Bothrops insularis (Island jararaca) (Queimada jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=90351557; PubMed=2386615;
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
RT "Primary structure and biological activity of bradykinin potentiating peptides from Bothrops insularis snake venom.";
RL J. Protein Chem. 9:221-227(1990).
CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
CC PIR; B37196; B37196.
DR PIR; B37196; B37196.
KW Hypotensive agent; Venom.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 10 AA; 1213 MW; 30C53546C761F773 CRC64;

Query Match 17.1%; Score 24; DB 1; Length 10;
Best Local Similarity 60.0%; Pred. No. 4.9e+02;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 11 HPPLP 15
| : |
Db 5 HPQIP 9

RESULT 7

ID MCA2_RHOOP STANDARD; PRT; 15 AA.
AC P56870;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE MALEYLACETATE REDUCTASE II (EC 1.3.1.32) (FRAGMENT).
OS Rhodococcus opacus (Nocardia opaca).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=37919;
RN [1]
RP SEQUENCE.
RC STRAIN=1CP;
RX MEDLINE=98324954; PubMed=9657989;
RA Seibert V., Kourbatova E.M., Golovleva L.A., Schloemann M.;
RT "Characterization of a maleylacetate reductase encoding region from

RT Rhodococcus opacus 1CP.";
RL J. Bacteriol. 180:3503-3508(1998).
CC -1- CATALYTIC ACTIVITY: 3-OXOADIPATE + NAD(P)(+) -> 2-MALEYLACETATE + NAD(P)H.
CC -1- PATHWAY: 3-CHLOROCATECHOL DEGRADATION (BETA-KETOADIPATE PATHWAY).
CC THIS PATHWAY SERVES A VITAL ROLE IN THE BIODEGRADATION OF TOXIC AROMATIC COMPOUNDS INTRODUCED IN THE ENVIRONMENT BOTH AS NATURAL PRODUCTS AND AS INDUSTRIAL EFFLUENT.
CC -1- SIMILARITY: BELONGS TO THE IRON-CONTAINING ALCOHOL DEHYDROGENASE FAMILY.
DR InterPro; IPR001670; Fe-ADH.
DR PROSITE; PS00913; ADH_IRON_1; PARTIAL.
DR PROSITE; PS00060; ADH_IRON_2; PARTIAL.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1884 MW; 58DA90DD038F025E CRC64;

Query Match 17.1%; Score 24; DB 1; Length 15;
Best Local Similarity 36.4%; Pred. No. 7.5e+02;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 8 RVTHPLRAL 18
| : | : | :
Db 2 RFEHNPQRI 12

RESULT 8

ID HI70_RAT STANDARD; PRT; 19 AA.
AC P21794;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DE HORMONE-INDUCED PROTEIN 70 KDA (HIP-70) (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RX MEDLINE=90208308; PubMed=2181662;
RA Mobbs C.V., Fink G., Pfaff D.W.;
RT "HIP-70: a protein induced by estrogen in the brain and LH-RH in the pituitary.";
RL Science 247:1477-1479(1990).
CC -1- INDUCTION: MOST PROMINENT PROTEIN INDUCED BY ESTROGEN IN HYPOTHALAMUS AND MOST PROMINENT PROTEIN INDUCED BY LH-RH IN PITUITARY.
CC NON_TER 19 19
SQ SEQUENCE 19 AA; 2113 MW; F793A98720B68E3C CRC64;

Query Match 17.1%; Score 24; DB 1; Length 19;
Best Local Similarity 50.0%; Pred. No. 9.7e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 ETYQSRVT 10
| : | : | :
Db 8 ENFESRVS 15

RESULT 9

ID CXAL_CONMA STANDARD; PRT; 14 AA.
AC P01521;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE ALPHA-CONOTOXIN MI (M1).
OS Conus magus (Magus cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda; Neogastropoda; Conoidea; Conidae; Conus.

OX NCBI_TaxID=6492;
RN [1]
RP SEQUENCE.
RX MEDLINE=83073458; PubMed=7149738;
RA McIntosh J.M., Cruz L.J., Hunkapiller M.W., Gray W.R., Olivera B.M.;
RT "Isolation and structure of a peptide toxin from the marine snail
RT Conus magus";
RL Arch. Biochem. Biophys. 218:329-334(1982).
RN [2]
RP DISULFIDE BONDS.
RX MEDLINE=84032400; PubMed=6630187;
RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;
RT "Conotoxin M1. Disulfide bonding and conformational states";
RL J. Biol. Chem. 258:12247-12251(1983).
RN [3]
RP REVIEW.
RX MEDLINE=89024586; PubMed=3052286;
RA Gray W.R., Olivera B.M., Cruz L.J.;
RT "Peptide toxins from venomous Conus snails";
RL Annu. Rev. Biochem. 57:665-700(1988).
CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
DR PIR; A01784; NTKNIM.
KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom.
FT DISULFID 3 8
FT DISULFID 4 14
FT MOD_RES 14 14 AMIDATION.
SQ SEQUENCE 14 AA; 1499 MW; DEE91898BF5E5BD CRC64;

Query Match 16.4%; Score 23; DB 1; Length 14;
Best Local Similarity 60.0%; Pred. No. 1e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEY 5
Db [1]
8 CGKNY 12

RESULT 10
UN05_PINPS STANDARD; PRT; 20 AA.
AC P81674;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF NEEDLES (N147) (FRAGMENTS).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RP SEQUENCE.
RC TISSUE=Needle;
RX MEDLINE=99274088; PubMed=10344291;
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
RT proteins";
RL Electrophoresis 20:1098-1108(1999).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.6, ITS MW IS: 36 KDA.
FT NON_TER 1 1
FT NON_CONS 11 12 F -> I.
FT VARIANT 13 13 Y -> E.
FT VARIANT 14 14 Y -> E.
FT VARIANT 15 15 R -> K.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2438 MW; 9F4E4678E086C298 CRC64;

Query Match 16.4%; Score 23; DB 1; Length 20;
Best Local Similarity 80.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 12 PHLPR 16
Db [1]
16 PELPR 20

RESULT 11
FA71_TETPY STANDARD; PRT; 21 AA.
AC P81424;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 71 KDA F-ACTIN BINDING PROTEIN (FRAGMENT).
OS Tetrahymena pyriformis.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Hymenostomatida;
OC Tetrahymena; Tetrahymena.
OX NCBI_TaxID=5908;
RN [1]
RP SEQUENCE.
RX MEDLINE=98207047; PubMed=9538250;
RA Watanabe A., Kurasawa Y., Watanabe Y., Numata O.;
RT "A new Tetrahymena actin-binding protein is localized in the division
RT furrow";
RL J. Biochem. 123:607-613(1998).
CC -!- FUNCTION: BINDS DIRECTLY TO F-ACTIN AND INDUCES ACTIN FILAMENT
CC BUNDLING. MAY FUNCTION AS A REGULATOR OF ACTIN FILAMENT
CC ORGANIZATION.
CC -!- DEVELOPMENTAL STAGE: CO-LOCALIZED WITH ACTIN IN THE ORAL APPARATUS
CC IN INTERPHASE CELLS. IN DIVIDING CELLS CO-LOCALIZED WITH ACTIN IN
CC THE DIVISION FURROW.
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
CC 7.5, ITS MW IS: 71 KDA.
CC -!- SIMILARITY: TO YEAST FIMBRIN.
KW Actin-binding.
FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2691 MW; 104FD6FD5E08FD28 CRC64;

Query Match 16.4%; Score 23; DB 1; Length 21;
Best Local Similarity 66.7%; Pred. No. 1.5e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 HPHLPR 16
Db [1]
10 HPLYLR 15

RESULT 12
SODM_RANCA STANDARD; PRT; 23 AA.
AC P36215;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SUPEROXIDE DISMUTASE [MN], MITOCHONDRIAL (EC 1.15.1.1) (FRAGMENT).
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=87126854; PubMed=3492965;
RA Abe Y., Okazaki T.;
RT "Purification and properties of the manganese superoxide dismutase
RT from the liver of bullfrog, Rana catesbeiana";
RL Arch. Biochem. Biophys. 253:241-248(1987).
CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE

```

CC      CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC      -1- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC      -1- SUBUNIT: HOMOTETRAMER.
CC      -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC      -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC      FAMILY.
CC      HSPP; P04179; IMSD.
CC      InterPro: IPR001189; SOD_M1.
CC      Pfam: PF00081; sodfe; 1.
CC      PROSITE; PS00088; SOD_MN; PARTIAL.
CC      Oxidoreductase; Manganese; Mitochondrion.
CC      NON_TER 23
CC      SEQUENCE 23 AA; 2594 MW; 5D80ED9B0E04F625 CRC64;

Query Match 16.4%; Score 23; DB 1; Length 23;
Best Local Similarity 26.7%; Pred. No. 1.7e+03;
Matches 4; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 5 YQSRVTHPLPALM 19
   | | | |
Db 9 YDFGALQPHISAEIM 23

RESULT 13
BRL_RANBP
ID BRL_RANBP STANDARD; PRT; 24 AA.
AC P32423;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE BREVININ-1.
OS Rana brevipoda porsa (Frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX NCBI_TaxID=30358;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=93080555; PubMed=1449472;
RA Morikawa N., Hagiwara K., Nakajima T.;
RT "Brevinin-1 and -2, unique antimicrobial peptides from the skin of
the frog, Rana brevipoda porsa.";
RL Biochem. Biophys. Res. Commun. 189:184-190(1992).
CC -1- FUNCTION: SHOWS ANTIBACTERIAL ACTIVITY AGAINST REPRESENTATIVE
GRAM-NEGATIVE AND GRAM-POSITIVE BACTERIAL SPECIES, AND A VERY
HIGH HEMOLYTIC ACTIVITY.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: SKIN.
CC -1- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
FAMILY.
CC PIR; JCI355; JCI355.
CC Amphibian skin; Antibiotic; Hemolysis.
FT DISULFID 18 24 BY SIMILARITY.
SQ SEQUENCE 24 AA; 2531 MW; C866285B191EFD4 CRC64;

Query Match 16.4%; Score 23; DB 1; Length 24;
Best Local Similarity 33.3%; Pred. No. 1.7e+03;
Matches 4; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 14 LPRALMRSTTKC 25
   | | | | |
Db 13 VPALFCKITKCC 24

RESULT 14
GAE6_RANRU
ID GAE6_RANRU STANDARD; PRT; 24 AA.
AC P80400;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)

```

```

DE GAEGURIN-6.
OS Rana rugosa (Frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
OX NCBI_TaxID=8410;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin;
RX MEDLINE=95091844; PubMed=7999137;
RA Park J.M., Jung J.-E., Lee B.J.;
RT "Antimicrobial peptides from the skin of a Korean frog, Rana
rugosa.";
RL Biochem. Biophys. Res. Commun. 205:948-954(1994).
CC -1- FUNCTION: HAS A NON-HEMOLYTIC ACTIVITY. HAS A BROAD SPECTRUM OF
ACTIVITY AGAINST BOTH GRAM POSITIVE AND GRAM NEGATIVE BACTERIA,
FUNGI AND PROTOZOA.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- TISSUE SPECIFICITY: SKIN.
CC -1- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
FAMILY.
CC Amphibian skin; Antibiotic.
FT DISULFID 18 24 BY SIMILARITY.
SQ SEQUENCE 24 AA; 2610 MW; 09918123FF90CCFD CRC64;

Query Match 16.4%; Score 23; DB 1; Length 24;
Best Local Similarity 33.3%; Pred. No. 1.7e+03;
Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 14 LPRALMRSTTKC 25
   | | | | |
Db 13 LPTICKISYKC 24

RESULT 15
MALE_KLEPN
ID MALE_KLEPN STANDARD; PRT; 10 AA.
AC Q05564;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE MALTOSE-BINDING PERIPLASMIC PROTEIN (MALTODEXTRIN-BINDING PROTEIN)
(MMBP) (FRAGMENT).
GN MALE.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1033-5P14 / KAY2026;
RX MEDLINE=93211295; PubMed=8459773;
RA Bachellier S., Perrin D., Hofnung M., Gilson E.;
RT "Bacterial interspersed mosaic elements (BIMES) are present in the
genome of Klebsiella.";
RL Mol. Microbiol. 7:537-544(1993).
CC -1- FUNCTION: MALE IS INVOLVED IN THE HIGH-AFFINITY MALTOSE MEMBRANE
TRANSPORT SYSTEM. INITIAL RECEPTOR FOR THE ACTIVE TRANSPORT OF AND
CHEMOTAXIS TOWARD MALTOOLIGOSACCHARIDES.
CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
PROTEIN FAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X68329; CAA48406.1; -.

```

DR InterPro; IPR000567; SBP_bac_1.
DR PROSITE; PS01037; SBP_BACTERIAL_1; PARTIAL.
KW Transport; Sugar transport; Periplasmic.
FT NON_TER 1
SQ SEQUENCE 10 AA; 1159 MW; 8FD8DC4415A6DDDA CRC64;

Query Match 15.7%; Score 22; DB 1; Length 10;
Best Local Similarity 80.0%; Pred. NO. le+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 6 QSRVT 10
|||:
Db 5 QSRIT 9

Search completed: March 4, 2002, 13:20:28
Job time: 615 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:10:59 ; Search time 76.45 Seconds
(without alignments)
47.833 Million cell updates/sec

Title: US-09-701-623c-5

Perfect score: 140

Sequence: 1 GQYQSRVTHPLRALMRSTTKC 25

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 7775

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 33 | 23.6 | 21 | 4 Q9BU87 | Q9bu87 homo sapien |
| 2 | 31 | 22.1 | 17 | 2 Q34216 | Q34216 pseudomonas |
| 3 | 28 | 20.0 | 17 | 6 Q9TR98 | Q9tr98 canis famil |
| 4 | 28 | 20.0 | 23 | 2 Q95839 | Q95839 staphylococ |
| 5 | 27 | 19.3 | 25 | 4 Q9UBT0 | Q9ubt0 homo sapien |
| 6 | 26 | 18.6 | 18 | 4 Q13665 | Q13665 homo sapien |
| 7 | 26 | 18.6 | 18 | 4 Q9BQT0 | Q9bqt0 homo sapien |
| 8 | 26 | 18.6 | 20 | 6 Q9TQ26 | Q9tq26 macaca fasc |
| 9 | 26 | 18.6 | 23 | 2 Q52321 | Q52321 escherichia |
| 10 | 26 | 18.6 | 24 | 12 Q73337 | Q73337 human immun |
| 11 | 25.5 | 18.2 | 19 | 13 Q42416 | Q42416 gallus gall |
| 12 | 25.5 | 18.2 | 22 | 4 Q9BUH0 | Q9buh0 homo sapien |
| 13 | 25 | 17.9 | 17 | 4 Q16310 | Q16310 homo sapien |
| 14 | 25 | 17.9 | 19 | 4 Q16271 | Q16271 homo sapien |
| 15 | 25 | 17.9 | 19 | 4 Q9UCCL | Q9uccl homo sapien |
| 16 | 25 | 17.9 | 19 | 11 Q9R1F0 | Q9rlf0 mus musculu |
| 17 | 25 | 17.9 | 20 | 13 P82232 | P82232 rana tempor |
| 18 | 25 | 17.9 | 21 | 11 Q9QVT3 | Q9qvt3 mus sp. ret |
| 19 | 25 | 17.9 | 21 | 13 Q9PS29 | Q9ps29 bitis ariet |

| | | | | | |
|----|------|------|----|-----------|---------------------|
| 20 | 25 | 17.9 | 22 | 4 Q13726 | Q13726 homo sapien |
| 21 | 25 | 17.9 | 23 | 2 Q9R5C2 | Q9r5c2 nitrosomona |
| 22 | 25 | 17.9 | 23 | 8 Q99162 | Q99162 moema staec |
| 23 | 25 | 17.9 | 24 | 6 Q9TRX6 | Q9trx6 bos taurus |
| 24 | 25 | 17.9 | 24 | 13 P82871 | P82871 rana sylvat |
| 25 | 25 | 17.9 | 25 | 11 Q64341 | Q64341 mus musculu |
| 26 | 24.5 | 17.5 | 17 | 6 Q9XSG1 | Q9xsg1 bos taurus |
| 27 | 24 | 17.1 | 16 | 8 Q9TZR0 | Q9tzt0 solanum tub |
| 28 | 24 | 17.1 | 17 | 2 Q9APU5 | Q9apu5 pseudomonas |
| 29 | 24 | 17.1 | 19 | 11 Q9QV38 | Q9qv38 mus sp. erp |
| 30 | 24 | 17.1 | 19 | 12 Q9O625 | Q9o625 simian herp |
| 31 | 24 | 17.1 | 19 | 12 Q9W9U1 | Q9w9u1 simian herp |
| 32 | 24 | 17.1 | 20 | 2 Q50180 | Q50180 pseudomonas |
| 33 | 24 | 17.1 | 20 | 2 Q9R4F1 | Q9r4f1 desulfovibr |
| 34 | 24 | 17.1 | 20 | 5 Q46158 | Q46158 lumbricus r |
| 35 | 24 | 17.1 | 20 | 8 Q33294 | Q33294 zea mays (m |
| 36 | 24 | 17.1 | 20 | 12 Q9WLD2 | Q9wld2 simian herp |
| 37 | 24 | 17.1 | 21 | 2 Q9X3C4 | Q9x3c4 prochloroco |
| 38 | 24 | 17.1 | 22 | 4 Q13659 | Q13659 homo sapien |
| 39 | 24 | 17.1 | 23 | 5 Q94781 | Q94781 trypanosoma |
| 40 | 24 | 17.1 | 23 | 12 Q65291 | Q65291 human adeno |
| 41 | 24 | 17.1 | 25 | 5 Q94683 | Q94683 polyandroca |
| 42 | 24 | 17.1 | 25 | 5 Q9BLZ8 | Q9blz8 macrotrache |
| 43 | 24 | 17.1 | 25 | 5 Q9BLZ5 | Q9blz5 adineta vag |
| 44 | 24 | 17.1 | 25 | 5 Q9BH69 | Q9bh69 adineta vag |
| 45 | 24 | 17.1 | 25 | 6 Q77604 | Q77604 macaca mulla |

ALIGNMENTS

RESULT 1

Q9BU87 ID Q9BU87 PRELIMINARY; PRT; 21 AA.
AC Q9BU87;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE UNKNOWN (PROTEIN FOR MGC:3518).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ENDOMETRIAL ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002821; AAH02821.1; -
SQ SEQUENCE 21 AA; 2149 MW; EF66757E7B79C6EC CRC64;

Query Match 23.6%; Score 33; DB 4; Length 21;
Best Local Similarity 54.5%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 7 SRVTHPLPRA 17

Db 7 SKTRPEAPRA 17

RESULT 2

Q34216 ID Q34216 PRELIMINARY; PRT; 17 AA.
AC Q34216;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
DE 2-HYDROXYMUCONIC SEMIALDEHYDE DEHYDROGENASE (FRAGMENT).
GN PHNG.
OS Pseudomonas sp.
OC Bacteria; Proteobacteria.
OX NCBI_TaxID=306;

```

RN SEQUENCE FROM N.A.
RC STRAIN=DJ77;
RA MEDLINE=97445124; PubMed=9299451;
RX Kim S., Kwon O.K., Kim Y., Kim C.K., Lee K.S., Kim Y.C.;
RT "Localization and sequence analysis of the phnH gene encoding 2-
RL hydroxyphenyl-2,4-dienoate hydratase in Pseudomonas sp. strain DJ77.";
RT Biochem. Biophys. Res. Commun. 238:56-60(1997).
DR EMBL; U97697; AAB71650.1; -.
FT NON_TER 1
SQ SEQUENCE 17 AA; 2203 MW; 4E913B3E1212214B CRC64;

Query Match 22.1%; Score 31; DB 2; Length 17;
Best Local Similarity 60.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 11 HPPLPRALMR 20
Db 8 HQHLRKALRR 17

RESULT 3
Q9TR98 PRELIMINARY; PRT; 17 AA.
AC Q9TR98;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ALBUMIN (FRAGMENT).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RX MEDLINE=95007849; PubMed=7923441;
RA Miller M.J., Parrelee D.C., Benjamin T., Sechi S., Dooley K.L.,
RA Kadlubar F.F.;
RT "Plasma proteins as early biomarkers of exposure to carcinogenic
RL Chem.-Biol. Interact. 93:221-234(1994).
DR HSP; P02768; LUOR.
SQ SEQUENCE 17 AA; 2024 MW; 1D39F70F7D23B269 CRC64;

Query Match 20.0%; Score 28; DB 6; Length 17;
Best Local Similarity 44.4%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ETYQSRVTH 11
Db 1 EAYKSEIAH 9

RESULT 4
P95839 PRELIMINARY; PRT; 23 AA.
AC P95839;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE ORF56 (FRAGMENT).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=96427339; PubMed=8830703;
RA Wu S., de Lencastre H., Tomasz A.;
RT "Sigma-B, a putative operon encoding alternate sigma factor of
```

```

RT Staphylococcus aureus RNA polymerase: molecular cloning and DNA
RT sequencing.";
RL J. Bacteriol. 178:6036-6042(1996).
DR EMBL; Y09929; CAA71063.1; -.
FT NON_TER 1
SQ SEQUENCE 23 AA; 2541 MW; 7F47717B1767D34F CRC64;

Query Match 20.0%; Score 28; DB 2; Length 23;
Best Local Similarity 71.4%; Pred. No. 1.7e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGETYQS 7
Db 11 CNETYLS 17

RESULT 5
Q9UBT0 PRELIMINARY; PRT; 25 AA.
AC Q9UBT0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE 6-PHOSPHOFRUCTO-2-KINASE/FRUCTOSE-2, 6-BISPHOSPHATASE (EC 2.7.1.105)
DE (FRAGMENT).
GN PFK2 OR IPFK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Fukasawa M., Sakakibara R., Yamamoto N.;
RT "Identification of the Promoter Region of Human Placental 6-
RT Phosphofructo-2-kinase/Fructose-2,6-bisphosphatase Gene.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Mahlknecht U., Bucala R.;
RT "Homo sapiens Inducible 6-phosphofructo-2-kinase/fructose-2,6-
RT bisphosphatase (IPFK2), genomic sequence.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB033994; BAA89351.1; -.
DR EMBL; AF110958; AAD23989.1; -.
KW Kinase; Transferase.
FT NON_TER 25
SQ SEQUENCE 25 AA; 2984 MW; 4ACC24B43FDBB74C CRC64;

Query Match 19.3%; Score 27; DB 4; Length 25;
Best Local Similarity 40.9%; Pred. No. 2.6e+03;
Matches 9; Conservative 0; Mismatches 5; Indels 8; Gaps 1;

Qy 3 ETYQSRVT-----HPHLPR 16
Db 4 ELTQSRVQKIWVPVDRPSLPR 25

RESULT 6
Q13665 PRELIMINARY; PRT; 18 AA.
AC Q13665;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE CARDIAC MYOSIN BINDING PROTEIN (FRAGMENT).
GN MYBP-C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
```

RN
RP SEQUENCE FROM N.A.
RX MEDLINE=96083592; PubMed=7493025;
RA Watkins H., Conner D., Thierfelder L., Jarcho J.A., MacRae C.,
RA McKenna W.J., Maron B.J., Seidman J.G., Seidman C.E.;
RT "Mutations in the cardiac myosin binding protein-C gene on chromosome
RT 11 cause familial hypertrophic cardiomyopathy."
RL Nat. Genet. 11:434-437(1995).
DR EMBL: S80805; AAB35661.1; -.
FT NON_TER 1 1
SQ SEQUENCE 18 AA; 1891 MW; 3EC842CE85A19C97 CRC64;

Query Match 18.6%; Score 26; DB 4; Length 18;
Best Local Similarity 50.0%; Pred. No. 2.7e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYQSRVT 10
Db 9 CGGIYVCRAT 18

RESULT 7
Q9BQ70 PRELIMINARY; PRT; 18 AA.
ID Q9BQ70
AC Q9BQ70
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 2.2 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21195339; PubMed=11297743;
RA Holzmann K., Ambrosch I., Eibling L., Micksche M., Berger W.;
RT "A small upstream open reading frame causes inhibition of human major
RT vault protein expression from a ubiquitous mRNA splice variant."
RL FEBS Lett. 494:99-104(2001).
DR EMBL: AJ291367; CAC35315.1; -.
KW Hypothetical protein.
SQ SEQUENCE 18 AA; 2179 MW; 5D06F9A3F11CB828 CRC64;

Query Match 18.6%; Score 26; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PHLP 15
Db 11 PHLP 14

RESULT 8
Q9TQ26 PRELIMINARY; PRT; 20 AA.
ID Q9TQ26
AC Q9TQ26
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 14, Last annotation update)
DE ALBUMIN (FRAGMENT).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE.
RX MEDLINE=96273610; PubMed=8690030;
RA Nicolas M.G., Fujiki K., Murayama K., Suzuki M.T., Mineki R.,
RA Hayakawa M., Yoshikawa Y., Cho F., Kanai A.;

RT
RT "Studies on the mechanism of early onset macular degeneration in
RT cynomolgus (Macaca fascicularis) monkeys. I. Abnormal concentrations
RT of two proteins in the retina."
RL Exp. Eye Res. 62:211-219(1996).
SQ SEQUENCE 20 AA; 2411 MW; 5F1A6AEB5918F777 CRC64;

Query Match 18.6%; Score 26; DB 6; Length 20;
Best Local Similarity 44.4%; Pred. No. 3e+03;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 ETYQSRVTH 11
Db 11 DTHKSEVAH 19

RESULT 9
Q52321 PRELIMINARY; PRT; 23 AA.
ID Q52321
AC Q52321
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE PLASMID ECOR124 (FROM E.COLI) HSDS (FRAGMENT).
GN HSDS.
OS Escherichia coli.
OG Plasmid IncFIV R124.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89137712; PubMed=3265670;
RA Price C., Bickle T.A.;
RT "Evolution of DNA sequence specificity in type I restriction
RT enzymes."
RL Biochem. Soc. Trans. 16:942-943(1988).
DR EMBL: M27782; AAA56854.1; -.
KW Plasmid.
FT NON_TER 1 1
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2436 MW; 17351FFB990C1827 CRC64;

Query Match 18.6%; Score 26; DB 2; Length 23;
Best Local Similarity 54.5%; Pred. No. 3.5e+03;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 13 HLPALMRSTT 23
Db 3 HLPALMRSTT 13

RESULT 10
Q73337 PRELIMINARY; PRT; 24 AA.
ID Q73337
AC Q73337
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE C18MBC, COMPLETE SEQUENCE.
GN NEF.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HIV-1 C18MBC;
RX MEDLINE=96069819; PubMed=7481804;
RA Deacon N.J., Tsykin A., Solomon A., Smith K., Ludford-Menting M.,
RA Hooker D.J., McPhee D.A., Greenway A.L., Ellett A., Chatfield C.,
RA Lawson V.A., Crowe S., Maerz S., Sonza S., Learmont J., Sullivan J.S.,
RA Cunningham A., Dwyer D., Dowton D., Mills J.;

RT "Genomic structure of an attenuated quasi species of HIV-1 from a
 RT blood transfusion donor and recipients.";
 RL Science 270:988-991(1995).
 DR EMBL: U37270; AAC54551.1; -.
 SQ SEQUENCE 24 AA; 2760 MW; 10805DB2E4F7B73D CRC64;

Query Match 18.6%; Score 26; DB 12; Length 24;
 Best Local Similarity 35.7%; Pred. No. 3.6e+03;
 Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GETYQSRVTHPLP 15
 I: - - - - -
 Db 3 GKWSESSVVRHVP 16

RESULT 11
 O42416
 ID O42416 PRELIMINARY; PRT; 19 AA.
 AC O42416;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)
 DE THROMBOMUCIN (FRAGMENT).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McNagney K.M., Pettersson I., Rossi F., Flamme I., Shevchenko A.,
 RA Mann M., Graf T.,
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Y13977; CAI74310.1; -.
 FT NON_TER 1
 FT NON_TER 19
 SQ SEQUENCE 19 AA; 2181 MW; 483C3DF97E13EC19 CRC64;

Query Match 18.2%; Score 25.5; DB 13; Length 19;
 Best Local Similarity 42.9%; Pred. No. 3.4e+03;
 Matches 6; Conservative 2; Mismatches 1; Indels 5; Gaps 1;

Qy 11 HPPLP-----RALM 19
 I: |||||
 Db 3 HPDLPGFDGGRIL 16

RESULT 12
 Q9BUHO
 ID Q9BUHO PRELIMINARY; PRT; 22 AA.
 AC Q9BUHO;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE UNKNOWN (PROTEIN FOR MGC:3878).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ENDOMETRIAL ADENOCARCINOMA;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC002629; AAH02629.1; -.
 SQ SEQUENCE 22 AA; 2645 MW; 60B9F471AB613874 CRC64;

Query Match 18.2%; Score 25.5; DB 4; Length 22;
 Best Local Similarity 53.3%; Pred. No. 3.9e+03;
 Matches 8; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Qy 8 RVTH-PHLPRALMRS 21
 I: |||||
 Db 3 RVTWPHLAVOLQRN 17

RESULT 13
 Q16310
 ID Q16310 PRELIMINARY; PRT; 17 AA.
 AC Q16310;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-JAN-1999 (TReMBLrel. 09, Last sequence update)
 DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
 DE GC*2 PROTEIN (FRAGMENT).
 GN GC*2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95242701; PubMed=7725672;
 RA Kofler A., Braun A., Jenkins T., Serjeantson S.W., Cleve H.;
 RT "Characterization of mutants of the vitamin-D-binding protein/group
 RT specific component: GC aborigine (IA1) from Australian aborigines and
 RT South African blacks, and 2A9 from south Germany.";
 RL Vox Sang. 68:50-54(1995).
 DR EMBL: S77130; AAD14250.1; -.
 FT NON_TER 17
 FT NON_TER 17
 SQ SEQUENCE 17 AA; 1845 MW; BB26CAD60293722C CRC64;

Query Match 17.9%; Score 25; DB 4; Length 17;
 Best Local Similarity 36.4%; Pred. No. 3.7e+03;
 Matches 4; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 15 PRALMRSTTKC 25
 I: | : | |
 Db 6 PKELAKLVNKC 16

RESULT 14
 Q16271
 ID Q16271 PRELIMINARY; PRT; 19 AA.
 AC Q16271;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
 DE ARGININE VASOPRESSIN V2 RECEPTOR (FRAGMENT).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95086168; PubMed=7993996;
 RA Holtzman E.J., Kolakowski L.F. Jr., Gelfman-Holtzman O., O'Brien D.G.,
 RA Rasoulpour M., Guillot A.P., Austello D.A.;
 RT "Mutations in the vasopressin V2 receptor gene in two families with
 RT nephrogenic diabetes insipidus.";
 RL J. Am. Soc. Nephrol. 5:169-176(1994).
 DR EMBL: S75754; AAB32753.1; -.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 19 AA; 1905 MW; 181640EFD90F2788 CRC64;

Query Match 17.9%; Score 25; DB 4; Length 19;
 Best Local Similarity 80.0%; Pred. No. 4.1e+03;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 HPPLP 15
 I: |||
 Db 4 HPSLP 8

RESULT 15
Q9UCC1
ID Q9UCC1 PRELIMINARY; PRT; 19 AA.
AC Q9UCC1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE ECWASE (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=94081366; PubMed=8258716;
RA Froelich C.J., Zhang X., Turbov J., Hudig D., Winkler U., Hanna W.L.;
RT "Human granzyme B degrades aggrecan proteoglycan in matrix synthesized
RL J. Immunol. 151:7161-7171(1993).
SQ SEQUENCE 19 AA; 2172 MW; 9F6E339D1DB0E28 CRC64;

Query Match 17.9%; Score 25; DB 4; Length 19;
Best Local Similarity 80.0%; Pred. No. 4.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Caps 0;

QY 12 PHLPR 16
Db 9 PHSPR 13

Search completed: March 4, 2002, 13:21:51
Job time: 652 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:11:39 ; Search time 40.99 Seconds
(without alignments)
46.459 Million cell updates/sec

Title: US-09-701-623c-6
Perfect score: 141
Sequence: 1 CGETYYSRVTHPLPKDIVRSIAKC 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 4959

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|--------|--------------------|
| 1 | 29 | 20.6 | 20 | A61093 | glue protein - Cal |
| 2 | 28 | 19.9 | 11 | S78765 | ribosomal protein |
| 3 | 28 | 19.9 | 23 | F61491 | seed protein ws-6 |
| 4 | 27 | 19.1 | 15 | C37765 | hypothetical prote |
| 5 | 27 | 19.1 | 17 | PH1802 | T cell receptor al |
| 6 | 27 | 19.1 | 19 | B53145 | high conductance c |
| 7 | 27 | 19.1 | 19 | E49048 | T-cell receptor be |
| 8 | 27 | 19.1 | 20 | A41439 | acid ribonuclease |
| 9 | 26.5 | 18.8 | 25 | A58647 | alphaA-conotoxin p |
| 10 | 26 | 18.4 | 13 | S32551 | glutathione transf |
| 11 | 26 | 18.4 | 14 | PH1628 | Ig H chain V-D-J r |
| 12 | 26 | 18.4 | 15 | I38336 | hypothetical TEL/M |
| 13 | 26 | 18.4 | 16 | PH1634 | Ig H chain V-D-J r |
| 14 | 26 | 18.4 | 25 | A60621 | somatotropin - Atl |
| 15 | 25 | 17.7 | 10 | H37196 | bradykinin-potenti |
| 16 | 25 | 17.7 | 14 | S62374 | alpha-1-antichymot |
| 17 | 25 | 17.7 | 15 | PA0071 | superoxide dismuta |
| 18 | 25 | 17.7 | 15 | I50503 | agrin - electric r |
| 19 | 25 | 17.7 | 17 | S69164 | ferredoxin al - Ja |
| 20 | 25 | 17.7 | 21 | B33600 | glutamate--ammonia |
| 21 | 25 | 17.7 | 22 | PH1359 | Ig heavy chain DJ |
| 22 | 25 | 17.7 | 22 | T01859 | CyI la protein - c |
| 23 | 25 | 17.7 | 23 | S72535 | probable acr-2 reg |
| 24 | 25 | 17.7 | 24 | B44379 | omega-conotoxin SV |
| 25 | 25 | 17.7 | 24 | S29749 | serum albumin - do |
| 26 | 25 | 17.7 | 24 | G85602 | hypothetical prote |
| 27 | 25 | 17.7 | 24 | H85653 | hypothetical prote |
| 28 | 25 | 17.7 | 25 | PH1686 | Ig heavy chain V r |
| 29 | 24.5 | 17.4 | 23 | S23637 | hypothetical prote |

| | | | | | | |
|----|------|------|----|---|--------|---------------------|
| 30 | 24 | 17.0 | 10 | 2 | B37196 | bradykinin-potenti |
| 31 | 24 | 17.0 | 13 | 2 | S47358 | T-cell antigen rec |
| 32 | 24 | 17.0 | 15 | 2 | S42741 | ubiquinol--cytochr |
| 33 | 24 | 17.0 | 17 | 2 | I49425 | mitogen regulated |
| 34 | 24 | 17.0 | 20 | 2 | S23981 | outer layer protei |
| 35 | 24 | 17.0 | 22 | 2 | S42567 | cytochrome-b5 redu |
| 36 | 24 | 17.0 | 23 | 2 | S48156 | alpha-amyliase inhi |
| 37 | 24 | 17.0 | 24 | 2 | A25738 | nicotinic acetylch |
| 38 | 24 | 17.0 | 25 | 2 | S03456 | T-cell receptor al |
| 39 | 23.5 | 16.7 | 14 | 2 | A58963 | alpha-conotoxin Cn |
| 40 | 23 | 16.3 | 11 | 2 | PT0301 | Ig heavy chain CRD |
| 41 | 23 | 16.3 | 14 | 1 | NTKNIM | alpha-conotoxin MI |
| 42 | 23 | 16.3 | 15 | 2 | PA0009 | seed storage prote |
| 43 | 23 | 16.3 | 17 | 2 | C85956 | hypothetical prote |
| 44 | 23 | 16.3 | 19 | 2 | PA0012 | superoxide dismuta |
| 45 | 23 | 16.3 | 20 | 2 | JA0142 | proteinase inhibit |

ALIGNMENTS

RESULT 1
A61093
glue protein - California mussel (fragments)
N;Alternate names: adhesive polyphenolic protein
C;Species: Mytilus californianus (California mussel)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
C;Accession: A61093
R;Waite, J. H.
J. Comp. Physiol. B 156, 491-496, 1986
A;Title: Mussel glue from Mytilus californianus Conrad: a comparative study.
A;Reference number: A61093; MUID:86279063
A;Accession: A61093
A;Molecule type: protein
A;Residues: 1-20 <WAI>
A;Note: the amino terminal residue was tentatively identified as Ser
A;Note: 12-Ser was also found
C;Comment: This glue protein is a component of the adhesive plaque of the byssus, the

C;Keywords: hydroxyproline; tandem repeat
F;4.14/Modified site: 3-hydroxyproline (Pro) #status experimental
F;5.9,15,19/Modified site: 4-hydroxyproline (Pro) #status experimental
F;7.17/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental
F;13/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) (partial) #status experimental
Query Match 20.6% Score 29; DB 2; Length 20;
Best Local Similarity 40.0%; Pred. No. 6.3e+02;
Matches 6; Conservative 4; Mismatches 3; Indels 2; Gaps 1;
QY 4 TYYSRVTHP--HLPK 16
|| :||:| :||
DB 6 TYKPKITYPPYKPK 20

RESULT 2
S78765
ribosomal protein MRP-S24, mitochondrial - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: S78765
R;Graack, H.R.
submitted to the Protein Sequence Database, July 1999
A;Reference number: S78760
A;Accession: S78765
A;Molecule type: protein
A;Residues: 1-11 <GRA>
C;Keywords: mitochondrial
F;1-11/Product: ribosomal protein MRP-S24 (fragment) #status experimental <MAT>
Query Match 19.9%; Score 28; DB 2; Length 11;
Best Local Similarity 40.0%; Pred. No. 4.7e+02;

Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 11 HPPLPKDIVR 20

Db 2 HVDVPLDLTK 11

RESULT 3

F61491

seed protein ws-6 - winged bean (fragment)

C:Species: Psophocarpus tetragonolobus (winged bean)

C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 03-May-1996

C:Accession: F61491

R:Hirano, H.

J. Protein Chem. 8, 115-130, 1989

A:Title: Microsequence analysis of winged bean seed proteins electrophoretically separated from two-dimensional gels

A:Reference number: A61491; MUID:89351606

A:Accession: F61491

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-23 <HIR>

C:Superfamily: pathogenesis-related protein

C:Keywords: seed

Query Match

Best Local Similarity 19.9%; Score 28; DB 2; Length 23;

Matches 6; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 4 TYYSRVTHPLPKDIVRSI 22

Db 4 TYEDETTPVAPAILYKAI 22

RESULT 4

C37765

hypothetical protein (csma 3' region) - Chloroflexus aurantiacus (fragment)

C:Species: Chloroflexus aurantiacus

C>Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 30-Sep-1993

C:Accession: C37765

R:Theroux, S.J.; Redlinger, T.E.; Fuller, R.C.; Robinson, S.J.

J. Bacteriol. 172, 4497-4504, 1990

A:Title: Gene encoding the 5.7-kilodalton chlorosome protein of Chloroflexus aurantiacus

A:Reference number: A37765; MUID:90330558

A:Accession: C37765

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-15 <THE>

A:Cross-references: GB:M33964

Query Match

Best Local Similarity 19.1%; Score 27; DB 2; Length 15;

Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 YYSRVTHP 12

Db 6 YYSHTVTP 13

RESULT 5

PH1802

T cell receptor alpha chain V region (clone 3PBL V alpha 24-8) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999

C:Accession: PH1802

R:Porcellini, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.

J. Exp. Med. 178, 1-16, 1993

A:Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blood

A:Reference number: PH1754; MUID:93301585

A:Accession: PH1802

A>Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-17 <POR>

Query Match

Best Local Similarity 19.1%; Score 27; DB 2; Length 17;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGETYS 7

Db 3 CGERPYS 9

RESULT 6

B53145

high conductance calcium-activated potassium channel, maxi-K channel - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994

C:Accession: B53145

R:Knaus, H.G.; Garcia-Calvo, M.; Kaczorowski, G.J.; Garcia, M.L.

J. Biol. Chem. 269, 3921-3924, 1994

A:Title: Subunit composition of the high conductance calcium-activated potassium chan

A:Reference number: A53145; MUID:94140798

A:Accession: B53145

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-19 <KNA>

A:Note: sequence extracted from NCBI backbone (NCBIP:144547)

Query Match

Best Local Similarity 19.1%; Score 27; DB 2; Length 19;

Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 3 EYYYSRVTHPH 13

Db 6 EYYQGSVLNPH 16

RESULT 7

E49048

T-cell receptor beta chain V region (CDR3 junction, clone Vbeta20.1) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997

C:Accession: E49048

R:Sioud, M.; Kjeldsen-Kragh, J.; Suleyman, S.; Vinje, O.; Natvig, J.B.; Forre, O.

Eur. J. Immunol. 22, 2413-2418, 1992

A:Title: Limited heterogeneity of T cell receptor variable region gene usage in juven

A:Reference number: A49048; MUID:92387250

A:Accession: E49048

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-19 <SIO>

A:Experimental source: patient EV, IL-2R+ synovial T-cells

A:Note: sequence extracted from NCBI backbone (NCBIP:113267)

C:Keywords: T-cell receptor

Query Match

Best Local Similarity 19.1%; Score 27; DB 2; Length 19;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGETYS 6

Db 9 CGRPYS 14

RESULT 8

A41439

acid ribonuclease (EC 3.1.1.1) - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 21-Mar-1996

C:Accession: A41439

R;Ohgi, K.; Sanda, A.; Takizawa, Y.; Irie, M.
J. Biochem. 103, 267-273, 1988
A:Title: Purification of acid ribonucleases from bovine spleen.
A:Reference number: A41439; MUID:88227899
A:Accession: A41439
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <OHG>
C:Keywords: hydrolase

Query Match 19.1%; Score 27; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 HPHPKDI 18
| : |||
Db 9 HLYFPKDL 16

RESULT 9

A58647
alphaA-conotoxin PIVA [validated] - cone shell (Conus purpurascens)
C:Species: Conus purpurascens (purple cone)
C:Date: 31-Oct-1997 #sequence_revision 07-Nov-1997 #text_change 15-Sep-2000
C:Accession: A58647
R;Hopkins, C.; Grille, M.; Miller, C.; Shon, K.J.; Cruz, L.J.; Gray, W.R.; Dykert, J.;
J. Biol. Chem. 270, 22361-22367, 1995
A:Title: A new family of Conus peptides targeted to the nicotinic acetylcholine receptor
A:Reference number: A58647; MUID:95403432
A:Accession: A58647
A:Molecule type: protein
A:Residues: 1-25 <HOP>
R;Han, K.H.; Hwang, K.J.; Kim, S.M.; Kim, S.K.; Gray, W.R.; Olivera, B.M.; Rivier, J.; S
submitted to the Brookhaven Protein Data Bank, December 1996
A:Reference number: A67666; PDB:1p1p
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue
F;7,13/Modified site: 4-hydroxyproline (Pro) #status experimental
Biochemistry 36, 1669-1677, 1997
A:Title: NMR structure determination of a novel conotoxin, [Pro 7,13] alpha A-conotoxin
A:Reference number: A58646; MUID:97200721
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
C:Superfamily: unassigned conotoxins
C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; hydroxyproline; pos
F;2-16,3-11,14-23/Disulfide bonds: #status experimental
F;7,13/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F;20/Modified site: 4-hydroxyproline (Pro) #status experimental
F;25/Modified site: amidated carboxyl end (Gln) #status experimental

Query Match 18.8%; Score 26.5; DB 2; Length 25;
Best Local Similarity 41.2%; Pred. No. 1.8e+03;
Matches 7; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 1 CGETYSRVTHPHLPKD 17
| : | : ||
Db 3 CG-SYPNACHPCSKD 18

RESULT 10

S32551
glutathione transferase (EC 2.5.1.18) mu (isoform pI 6.4) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Dec-1993 #sequence_revision 01-Dec-1995 #text_change 05-Jan-1996
C:Accession: S32551; S32550
R;Singhal, S.S.; Saxena, M.; Ahmad, H.; Awasthi, Y.C.
Biochim. Biophys. Acta 1116, 137-146, 1992
A:Title: Glutathione S-transferases of mouse liver: sex-related differences in the exper
A:Reference number: S32548; MUID:92256466
A:Accession: S32551
A:Molecule type: protein
A:Residues: 1-13 <SINI>
A:Experimental source: female

A:Accession: S32550
A:Molecule type: protein
A:Residues: 1-13 <SIN2>
A:Experimental source: male
C:Keywords: transferase

Query Match 18.4%; Score 26; DB 2; Length 13;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 8 RVTHP 12
| : |||
Db 9 RLTHP 13

RESULT 11

PH1628
Ig H chain V-D-J region (clone B-less 151) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C:Accession: PH1628
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less m
A:Reference number: PH1580; MUID:93301609
A:Accession: PH1628
A:Molecule type: DNA
A:Residues: 1-14 <LFV>
A:Experimental source: bone marrow pre-B lymphocyte
C:Keywords: immunoglobulin

Query Match 18.4%; Score 26; DB 2; Length 14;
Best Local Similarity 44.4%; Pred. No. 1.2e+03;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRV 9
| : |||
Db 1 CARYHSNL 9

RESULT 12

I38336
hypothetical TEL/MN1 mutant fusion protein type I - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 20-Apr-2000
C:Accession: I38336
R;Buijs, A.; Sherr, S.; van Baal, S.; van Bezouw, S.; van der Plas, D.; Van Kessel, A
Oncogene 10, 1511-1519, 1995
A:Title: Translocation (12;22) (p13;q11) in myeloproliferative disorders results in f
A:Reference number: I38031; MUID:95249265
A:Accession: I38336
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-15 <BUI>
A:Cross-references: EMBL:X85026; MUID:9971473; PIDN:CAA59399.1; PID:9971474
C:Comment: This sequence is the chimeric product of a translocation mutation.
C:Genetics:

A:Gene: ETV6/MN1; TEL/MN1
A:Map position: 22q11/12p13
C:Keywords: fusion protein

Query Match 18.4%; Score 26; DB 4; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 HLPKDI 18
| : |||
Db 6 HLPKDL 11

RESULT 13

PH1634
Ig H chain V-D-J region (clone B-less 213) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C:Accession: PH1634
R:Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A:Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A:Reference number: FH1580; MUID:93301609
A:Accession: PH1634
A:Molecule type: DNA
A:Residues: 1-16 <LEV>
A:Experimental source: bone marrow pre-B lymphocyte
C:Keywords: immunoglobulin

Query Match 18.4%; Score 26; DB 2; Length 16;

Best Local Similarity 41.7%; Pred. No. 1.4e+03;
Matches 5; Conservative 3; Mismatches 0; Indels 4; Gaps 1;

Qy 3 EYYXSRVTHPHL 14

:::|::|::|

Db 4 DSYYS----PHI 11

RESULT 14

A60621
Somatotropin - Atlantic salmon (fragment)
N:Alternate names: growth hormone
C:Species: Salmo salar (Atlantic salmon)
C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 30-Sep-1993
C:Accession: A60621
R:Skibeli, V.; Andersen, O.; Gautvik, K.M.
Gen. Comp. Endocrinol. 80, 333-344, 1990
A:Title: Purification and characterization of Atlantic salmon growth hormone and evidence of its biological activity
A:Reference number: A60621; MUID:91146880
A:Accession: A60621
A:Molecule type: protein
A:Residues: 1-25 <SKI>
A:Note: this protein displayed charge heterogeneity from variable levels of phosphorylation
C:Superfamily: prolactin
C:Keywords: hormone; phosphoprotein; pituitary

Query Match

18.4%; Score 26; DB 2; Length 25;

Best Local Similarity 62.5%; Pred. No. 2.2e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 SRVTHPHL 14

:|:|:|

Db 12 NRVOHLHL 19

RESULT 15

H37196
bradykinin-potentiating peptide 8 - island jararaca
C:Species: Bothrops insularis (island jararaca)
C:Date: 14-Feb-1992 #sequence_revision 01-Dec-1992 #text_change 05-Aug-1994
C:Accession: H37196
R:Cintra, A.C.O.; Vieira, C.A.; Giglio, J.R.
J. Protein Chem. 9, 221-227, 1990
A:Title: Primary structure and biological activity of bradykinin potentiating peptides from Bothrops insularis
A:Reference number: A37196; MUID:90351557
A:Accession: H37196
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-10 <CIN>
C:Keywords: pyroglutamic acid
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match

17.7%; Score 25; DB 2; Length 10;

THIS PAGE BLANK (USP 10)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:20:28 ; Search time 24.21 Seconds
(without alignments)
37.861 Million cell updates/sec

Title: US-09-701-623C-6
Perfect score: 141
Sequence: 1 CGETYYSRVTHPLPKDIVRSIAK 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1446

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 28 | 19.9 | 15 | 1 | MAOX_CHICK |
| 2 | 26.5 | 18.8 | 25 | 1 | CXA4_CONPU |
| 3 | 26 | 18.4 | 20 | 1 | UN05_PINPS |
| 4 | 25 | 17.7 | 10 | 1 | BPP8_BOTIN |
| 5 | 25 | 17.7 | 15 | 1 | MA02_RHOOP |
| 6 | 25 | 17.7 | 24 | 1 | CXA0A_CONST |
| 7 | 24 | 17.0 | 10 | 1 | BPP2_BOTIN |
| 8 | 24 | 17.0 | 10 | 1 | VM02_CHICK |
| 9 | 23.5 | 16.7 | 14 | 1 | CXA1_CONCN |
| 10 | 23 | 16.3 | 14 | 1 | CXA1_CONMA |
| 11 | 23 | 16.3 | 20 | 1 | MI17_BOVIN |
| 12 | 23 | 16.3 | 20 | 1 | SODM_HORVU |
| 13 | 23 | 16.3 | 21 | 1 | CFPA_TREPH |
| 14 | 23 | 16.3 | 23 | 1 | SODM_RANCA |
| 15 | 23 | 16.3 | 24 | 1 | DMS6_PHYBI |
| 16 | 23 | 16.3 | 25 | 1 | AND1_ANDAU |
| 17 | 22 | 15.6 | 13 | 1 | UP71_LITEW |
| 18 | 22 | 15.6 | 15 | 1 | DCMW_PSECA |
| 19 | 22 | 15.6 | 16 | 1 | CT12_LITCI |
| 20 | 22 | 15.6 | 16 | 1 | CT13_LITCI |
| 21 | 22 | 15.6 | 16 | 1 | MMPX_SOLTU |
| 22 | 22 | 15.6 | 17 | 1 | BOL4_MEGPE |
| 23 | 22 | 15.6 | 18 | 1 | CT14_LITCI |
| 24 | 22 | 15.6 | 18 | 1 | CT1D_LITCI |
| 25 | 22 | 15.6 | 21 | 1 | TERT_APIME |
| 26 | 22 | 15.6 | 22 | 1 | YHV4_LACHE |
| 27 | 21 | 14.9 | 13 | 1 | SODM_CANFA |
| 28 | 21 | 14.9 | 13 | 1 | UP51_UPEIN |
| 29 | 21 | 14.9 | 14 | 1 | DCMW_PSECF |
| 30 | 21 | 14.9 | 15 | 1 | DCMW_PSECH |
| 31 | 21 | 14.9 | 16 | 1 | CT11_LITCI |
| 32 | 21 | 14.9 | 16 | 1 | YMOR_PSEPU |
| 33 | 21 | 14.9 | 17 | 1 | AP1D_BOWPA |

| | | | | | |
|----|----|------|----|---|------------|
| 34 | 21 | 14.9 | 18 | 1 | A2M_OCTVU |
| 35 | 21 | 14.9 | 18 | 1 | CT1A_LITCI |
| 36 | 21 | 14.9 | 18 | 1 | CT1B_LITCI |
| 37 | 21 | 14.9 | 19 | 1 | AL22_HORSE |
| 38 | 21 | 14.9 | 19 | 1 | CXA2_CONST |
| 39 | 21 | 14.9 | 20 | 1 | APAL_EYRPA |
| 40 | 21 | 14.9 | 21 | 1 | PA71_TETPY |
| 41 | 21 | 14.9 | 21 | 1 | RL5_HALME |
| 42 | 21 | 14.9 | 22 | 1 | FUC1_RAT |
| 43 | 21 | 14.9 | 22 | 1 | FUC2_RAT |
| 44 | 21 | 14.9 | 23 | 1 | RL5_HALHA |
| 45 | 21 | 14.9 | 24 | 1 | ACHA_ELEEL |

ALIGNMENTS

RESULT 1
MAOX_CHICK STANDARD; PRT; 15 AA.
ID MAOX_CHICK
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADP-DEPENDENT MALIC ENZYME (EC 1.1.1.40) (NADP-ME) (FRAGMENT).
GN MEI.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97056061; PubMed=8900406;
RA Hodnett D.W., Fantozzi D.A., Thurmond D.C., Klautsky S.A.,
RA Macphree K.G., Estrem S.T., Xu G., Goodridge A.G.;
RT "The chicken malic enzyme gene: structural organization and
RT identification of triiodothyronine response elements in the
RT 5'-flanking DNA";
RL Arch. Biochem. Biophys. 334:309-324(1996).
CC -!- CATALYTIC ACTIVITY: L-MALATE + NADP(+) -> PYRUVATE + CO(2) + NADPH.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO THE MALIC ENZYMES FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL: U49693; AAA9721.1;
DR InterPro: IPRO01891; Malic_enzyme.
DR PROSITE: PS00331; MALIC_ENZYMES; PARTIAL.
KW Oxidoreductase; NADP.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1842 MW; CFEF180B2BA84C2B CRC64;

Query Match 19.9%; Score 28; DB 1; Length 15;
Best Local Similarity 54.5%; Pred. No. 2,7e+02;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 6 YSRVTHPLPK 16
| | | | |
Db 5 YEVVRDPHLNR 15

RESULT 2
CXA4_CONPU STANDARD; PRT; 25 AA.
ID CXA4_CONPU

P55963;
01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
15-DEC-1998 (Rel. 37, Last annotation update)
ALPHA-A CONOTOXIN PIVA.
Conus purpurascens (Purple cone).
Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
Neogastropoda; Conoidea; Conidae; Conus.
NCBI_TaxID=41690;
[1]
SEQUENCE.
TISSUE=Venom;
MEDLINE=95403432; PubMed=7673220;
Hopkins C., Grille M., Miller C., Shon K.-J., Cruz L.J., Gray W.R.,
Dykert J., Rivier J., Yoshikami D., Olivera B.M.;
"A new family of Conus peptides targeted to the nicotinic
acetylcholine receptor.";
J. Biol. Chem. 270:22361-22367(1995).
[2]
STRUCTURE BY NMR.
MEDLINE=97200721; PubMed=9048550;
Han K.-H., Hwang K.-J., Kim S.-M., Kim S.-K., Gray W.R., Olivera B.M.,
Rivier J., Shon K.-J.;
"NMR structure determination of a novel conotoxin, [Pro 7,13] alpha
A-conotoxin PIVA.";
Biochemistry 36:1669-1677(1997).
-1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
INHIBIT THEM.
PDB; IPIP; 07-JUL-97.
Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
Hydroxylation; Venom; 3D-structure.
FT DISULFID 2 16
FT DISULFID 3 11
FT DISULFID 14 23
FT MOD_RES 7 7 HYDROXYLATION.
FT MOD_RES 13 13 HYDROXYLATION.
FT MOD_RES 20 20 HYDROXYLATION.
FT MOD_RES 25 25 AMIDATION
SQ SEQUENCE 25 AA; 2608 MW; 9E2147898D697640 CRC64;

Query Match 18.8%; Score 26.5; DB 1; Length 25;
Best Local Similarity 41.2%; Pred. No. 7.7e+02;
Matches 7; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 1 CGETYYSRVTHPLPKD 17
Db 3 CG-SYPNAACHPCSKD 18
[1] : : : : :
[2] : : : : :

RESULT 3
UN05_PINPS STANDARD; PRT; 20 AA.
AC P81674;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF NEEDLES (N147) (FRAGMENTS).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
SEQUENCE.
TISSUE=Needle;
MEDLINE=99274088; PubMed=10344291;
Costa P., Plomont C., Bauw G., Dubos C., Bahrman N., Kremer A.,
Erigerio J.-M., Plomont C.;
"Separation and characterization of needle and xylem maritime pine
proteins.";
Electrophoresis 20:1098-1108(1999).
CC -b- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN

CC PROTEIN IS: 5.6, ITS MW IS: 36 KDA.
FT NON_TER 1 1
FT NON_CONS 11 12 F -> I.
FT VARIANT 13 13 Y -> E.
FT VARIANT 14 14 Y -> E.
FT VARIANT 15 15 R -> K.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2438 MW; 9F4E4678E086C298 CRC64;

Query Match 18.4%; Score 26; DB 1; Length 20;
Best Local Similarity 31.2%; Pred. No. 7.2e+02;
Matches 5; Conservative 3; Mismatches 4; Indels 4; Gaps 1;

QY 1 CGETYYSRVTHPLPK 16
Db 9 CDKDFY----RPPLPR 20
[1] : : : : :
[2] : : : : :

RESULT 4
BPP8_BOTIN STANDARD; PRT; 10 AA.
AC P30426;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE BRADYKININ-POTENTIATING PEPTIDE S5.1 (ANGIOTENSIN-CONVERTING
ENZYME INHIBITOR).
OS Bothrops insularis (Island Jararaca) (Queimada Jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodonta; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=90351557; PubMed=2386615;
Cintra A.C.O., Vieira C.A., Giglio J.R.;
"Primary structure and biological activity of bradykinin potentiating
peptides from Bothrops insularis snake venom.";
J. Protein Chem. 9:221-227(1990).
-1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
CC PIR; H37196; H37196.
DR Hypotensive agent; Venom.
KW MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
FT MOD_RES 1 1
SQ SEQUENCE 10 AA; 1173 MW; 2FF835545761F6D8 CRC64;

Query Match 17.7%; Score 25; DB 1; Length 10;
Best Local Similarity 60.0%; Pred. No. 4.9e+02;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 HPPLP 15
Db 5 HPNIP 9
[1] : : : : :
[2] : : : : :

RESULT 5
MCA2_RHOOP STANDARD; PRT; 15 AA.
AC P56870;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE MALEYLACETATE REDUCTASE II (EC 1.3.1.32) (FRAGMENT).
OS Rhodococcus opacus (Nocardia opaca).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=37919;
RN [1]

RP SEQUENCE.
RC STRAIN-ICP;
RX MEDLINE-98324954; PubMed-9657989;
RA Seibert V., Kourbatova E.M., Golovleva L.A., Schloemann M.;
RT "Characterization of a maleylacetate reductase encoding region from
RT Rhodococcus opacus ICP.";
RL J. Bacteriol. 180:3503-3508(1998).
CC -1- CATALYTIC ACTIVITY: 3-OXOADIPATE + NAD(P)(+) -> 2-MALEYLACETATE +
CC NAD(P)H.
CC -1- PATHWAY: 3-CHLOROCATECHOL DEGRADATION (BETA-KETOADIPATE PATHWAY).
CC THIS PATHWAY SERVES A VITAL ROLE IN THE BIODEGRADATION OF TOXIC
CC AROMATIC COMPOUNDS INTRODUCED IN THE ENVIRONMENT BOTH AS NATURAL
CC PRODUCTS AND AS INDUSTRIAL EFFLUENT.
CC -1- SIMILARITY: BELONGS TO THE IRON-CONTAINING ALCOHOL DEHYDROGENASE
CC FAMILY.
DR InterPro: IPR001670; Fe-ADH.
DR PROSITE: PS00913; ADH_IRON_1; PARTIAL.
DR PROSITE: PS00060; ADH_IRON_2; PARTIAL.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1884 MW; 58DA90DD038F025E CRC64;

Query Match 17.7%; Score 25; DB 1; Length 15;
Best Local Similarity 45.5%; Pred. No. 7.5e+02;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 RVTHPLPKDI 18
| | : | : |
DB 2 RFEHENLPQRI 12

RESULT 6

CXOA_CONST STANDARD; PRT; 24 AA.
ID AC P28860;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE OMEGA-CONOTOXIN SVIA.
OS Conus striatus (Striated cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6493;
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=Venom;
RX MEDLINE-93003172; PubMed-1390774;
RA Ramilo C., Zafaralla G.C., Nadasdi L., Hammerland L.G., Yoshikami D.,
RA Gray W.R., Kristipati R., Ramachandran J., Miljanich G., Olivera B.M.,
RA Cruz L.J.;
RT "Novel alpha- and omega-conotoxins from Conus striatus venom.";
RL Biochemistry 31:9919-9926(1992).
CC -1- FUNCTION: OMEGA-CONOTOXINS ACT AT PRESYNAPTIC MEMBRANES, THEY BIND
CC AND BLOCK THE CALCIUM CHANNELS.
DR PIR; B44379; B44379.
KW Presynaptic neurotoxin; Calcium channel inhibitor; Venom;
KW Amidation; Hydroxylation.
FT DISULFID 1 15
FT DISULFID 8 18
FT DISULFID 14 23
FT MOD_RES 7 7 HYDROXYLATION.
FT MOD_RES 24 24 AMIDATION.
SQ SEQUENCE 24 AA; 2485 MW; B29EFC982AB8E644 CRC64;

Query Match 17.7%; Score 25; DB 1; Length 24;
Best Local Similarity 40.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEYYSRV 10
| | : | : |
DB 15 CGRCYRGKCT 24

RESULT 7
BPP2_BOTIN STANDARD; PRT; 10 AA.
ID AC P30422;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE BRADYKININ-POTENTIATING PEPTIDE S4,3,1 (10C) (ANGIOTENSIN-CONVERTING
DE ENZYME INHIBITOR).
OS Bothrops insularis (Island Jararaca) (Queimada Jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE-90351557; PubMed-2386615;
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
RT "Primary structure and biological activity of bradykinin potentiating
RT peptides from Bothrops insularis snake venom.";
RL J. Protein Chem. 9:221-227(1990).
CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
DR PIR; B37196; B37196.
KW Hypotensive agent; Venom.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 10 AA; 1213 MW; 30C53546C761F773 CRC64;

Query Match 17.0%; Score 24; DB 1; Length 10;
Best Local Similarity 60.0%; Pred. No. 6.9e+02;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 HPHP 15
| | : | : |
DB 5 HPQIP 9

RESULT 8

VMO2_CHICK STANDARD; PRT; 20 AA.
ID AC Q9PS49;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE VITELLINE MEMBRANE OUTER LAYER PROTEIN II (VMO-II) (FRAGMENT).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RC TISSUE=Egg yolk;
RX MEDLINE-92392273; PubMed-1520265;
RA Kido S., Morimoto A., Kim F., Doi Y.;
RT "Isolation of a novel protein from the outer layer of the vitelline
RT membrane.";
RL Biochem. J. 286:17-22(1992).
CC -1- FUNCTION: EXACT FUNCTION NOT KNOWN. COMPONENT OF THE OUTER
CC MEMBRANE OF THE VITELLINE LAYER OF THE EGG.
CC -1- PTM: ALL CYSTEINE RESIDUES OF THE MATURE PROTEIN ARE INVOLVED IN
CC DISULFIDE BONDS.
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2325 MW; 45FC7989AB7527C7 CRC64;

Query Match 17.0%; Score 24; DB 1; Length 20;

Best Local Similarity 44.4%; Pred. No. 1.4e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 14 LPKDIVERSI 22
|:|:|:|:
Db 1 LPRDTSRXV 9

RESULT 9
CXAL_CONCN STANDARD; PRT; 14 AA.
AC P36973;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ALPHA-CONOTOXIN CNIA [CONTAINS: ALPHA-CONOTOXIN CNIB].
OS Conus consors.
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=101297;
RN [1]
RP SEQUENCE, SYNTHESIS, CHARACTERIZATION, AND STRUCTURE BY NMR.
RC TISSUE-Venom;
RX MEDLINE=99253590; PubMed=10320362;
RA Favreau P., Krimm I., le Gall F., Bobenrieth M.J., Lamthanh H.,
RA Bouet F., Servent D., Molgo J., Menez A., Letourneux Y.,
RA Lancelin J.-M.;
RT "Biochemical characterization and nuclear magnetic resonance
structure of novel alpha-conotoxins isolated from the venom of Conus
consors.";
RL Biochemistry 38:5317-5326(1999).
CC -|- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
INHIBIT THEM. THIS PEPTIDE SEEMS TO BE A POTENT AND SELECTIVE
BLOCKER OF MUSCULAR SUBTYPE OF NACHR.
CC PDB; 1B45; 09-JUL-99.
DR Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom; 3D-structure.
FT PEPTIDE 1 14 ALPHA-CONOTOXIN CNIA.
FT PEPTIDE 3 14 ALPHA-CONOTOXIN CNIB.
FT DISULFID 3 8
FT DISULFID 4 14
FT MOD_RES 14 14
SQ SEQUENCE 14 AA; 1548 MW; DEEE91969BF5E5BD CRC64;

Query Match 16.7%; Score 23.5; DB 1; Length 14;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 CGETYYS 7
|:|:|:|:
Db 8 CGK-YYS 13

RESULT 10
CXAL_CONMA STANDARD; PRT; 14 AA.
AC P01521;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE ALPHA-CONOTOXIN HI (MI).
OS Conus magus (Magus cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6492;
RN [1]
RP SEQUENCE

RX MEDLINE=83073458; PubMed=7149738;
RA McIntosh J.M., Cruz L.J., Hunkapiller M.W., Gray W.R., Olivera B.M.;
RT "Isolation and structure of a peptide toxin from the marine snail
Conus magus.";

RL Arch. Biochem. Biophys. 218:329-334(1982).
RN [2]
RP DISULFIDE BONDS.
RX MEDLINE=84032400; PubMed=6630187;
RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;
RT "Conotoxin MI. Disulfide bonding and conformational states.";
RL J. Biol. Chem. 258:12247-12251(1983).
RN [3]
RP REVIEW

RX MEDLINE=89024586; PubMed=3052286;
RA Gray W.R., Olivera B.M., Cruz L.J.;
RT "Peptide toxins from venomous Conus snails.";
ANNU. REV. BIOCHEM. 57:665-700(1988).
CC -|- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
INHIBIT THEM.
CC PIR; A01784; NTKNLM.
DR Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom.
FT DISULFID 3 8
FT DISULFID 4 14
FT MOD_RES 14 14
SQ SEQUENCE 14 AA; 1499 MW; DEEE91898BF5E5BD CRC64;

Query Match 16.3%; Score 23; DB 1; Length 14;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGETY 5
|:|:|:
Db 8 CGKNV 12

RESULT 11
MI17_BOVIN STANDARD; PRT; 20 AA.
ID MI17_BOVIN
AC P35451;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE 17 KDA MILK GLYCOPROTEIN (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE

RX TISSUE-Milk;
RX MEDLINE=93308294; PubMed=8320368;
RA Soerensen E.S., Petersen T.E.;
RT "Purification and characterization of three proteins isolated from
the proteose peptone fraction of bovine milk.";
RL J. Dairy Res. 60:189-197(1993).
CC -|- PTM: N-GLYCOSYLATED.
CC -|- SIMILARITY: TO CAMEL WHEY PROTEIN.
KW Glycoprotein; Milk.
FT NON_TER 1 1
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2233 MW; 4CCA589404C62C27 CRC64;

Query Match 16.3%; Score 23; DB 1; Length 20;
Best Local Similarity 40.0%; Pred. No. 2e+03;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 11 HPPLPKDIVR 20
|:|:|:|:
Db 9 NPKLPLSILK 18

RESULT 12

SODM_HORVU STANDARD; PRT; 20 AA.
ID F28524;
AC
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SUPEROXIDE DISMUTASE [MN], MITOCHONDRIAL (EC 1.15.1.1) (FRAGMENT).
GN SODA.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE.
RC STRAIN=CV, CM 72; TISSUE=Root;
RA Harkman W.J., Tao H.P., Tanaka C.K.;
RT "Germin-like polypeptides increase in barley roots during salt
stress";
RL Plant Physiol. 97:366-374(1991).
CC -I- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -I- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -I- SUBUNIT: HOMOTETRAMER.
CC -I- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC -I- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
DR HSSP; P04179; IMSD.
DR Mendel; 16446; HORVU:Soda:mn16446.
DR InterPro: IPR001189; SOD_MI.
DR Pfam: PF00081; sodfe; 1.
DR PROSITE; PS00088; SOD_MN; PARTIAL.
DR Oxidoreductase; Manganese; Mitochondrion.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2152 MW; 7F7CAB8DFF1C9657 CRC64;

Query Match 16.3%; Score 23; DB 1; Length 20;
Best Local Similarity 62.5%; Pred. No. 2e+03;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 10 THPHLPKD 17
Db 5 TLPDLPYD 12

RESULT 13
CFPA_TREPH STANDARD; PRT; 21 AA.
ID CFPA_TREPH
AC P56738;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOPLASMIC FILAMENT PROTEIN A (FRAGMENT).
GN CFPA.
OS Treponema phagedenis.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=162;
RN [1]
RP SEQUENCE.
RC STRAIN=KAZAN 5;
RA MEDLINE=96236033; PubMed=8655496;
RX You Y., Elmore S., Colton L.L., Mackenzie C., Stoops J.K.,
RA Weinstock G.M., Norris S.J.;
RT "Characterization of the cytoplasmic filament protein gene (cfpa) of
Treponema pallidum subsp. pallidum";
RL J. Bacteriol. 178:3177-3187(1996).
CC -I- FUNCTION: COMPONENT OF THE CYTOPLASMIC FILAMENTS THAT RUN THE
CC LENGTH OF THE ORGANISM JUST UNDERNEATH THE CYTOPLASMIC MEMBRANE.
CC -I- SUBCELLULAR LOCATION: AN ARRAY OF 4 TO 6 FILAMENTS LIE IN CLOSE
CC APPosition TO THE INNER MEMBRANE AND ARE ALWAYS LOCALIZED DIRECTLY
CC UNDERNEATH THE CORRESPONDING GROUP OF PERIPLASMIC FLAGELLA.
KW Structural protein; Antigen.

FT NON_TER 21 21
SQ SEQUENCE 21 AA; 2231 MW; 574604B4FFC2D017 CRC64;

Query Match 16.3%; Score 23; DB 1; Length 21;
Best Local Similarity 40.0%; Pred. No. 2.1e+03;
Matches 4; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 9 VTHPHLPKDI 18
Db 11 VFHPEKPSAV 20

RESULT 14
SODM_RANCA STANDARD; PRT; 23 AA.
ID SODM_RANCA
AC P36215;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SUPEROXIDE DISMUTASE [MN], MITOCHONDRIAL (EC 1.15.1.1) (FRAGMENT).
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=87126854; PubMed=3492965;
RA Abe Y., Okazaki T.;
RT "Purification and properties of the manganese superoxide dismutase
from the liver of bullfrog, Rana catesbeiana";
RL Arch. Biochem. Biophys. 253:241-248(1987).
CC -I- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -I- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -I- SUBUNIT: HOMOTETRAMER.
CC -I- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC -I- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
DR HSSP; P04179; IMSD.
DR InterPro: IPR001189; SOD_MI.
DR Pfam: PF00081; sodfe; 1.
DR PROSITE; PS00088; SOD_MN; PARTIAL.
DR Oxidoreductase; Manganese; Mitochondrion.
FT NON_TER 23
SQ SEQUENCE 23 AA; 2594 MW; 5D80ED9B0E04F625 CRC64;

Query Match 16.3%; Score 23; DB 1; Length 23;
Best Local Similarity 62.5%; Pred. No. 2.3e+03;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 10 THPHLPKD 17
Db 3 TLPDLPYD 10

RESULT 15
DMS6_PHYBI STANDARD; PRT; 24 AA.
ID DMS6_PHYBI
AC P81490;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DERMASEPTIN BVI (DERMASEPTIN B6).
OS Phyllomedusa bicolor (two-colored leaf frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OX NCBI_TaxID=8393;
RN [1]
RP SEQUENCE.

RC TISSUE-Skin;
RX MEDLINE-98278974; PubMed-9614066;
RA Charpentier S., Amiche M., Mester J., Vouille V., Le Caer J.-P.,
RA Nicolas P., Delfour A.;
RT "Structure, synthesis, and molecular cloning of dermaseptins B, a
family of skin peptide antibiotics.";
RL J. Biol. Chem. 273:14690-14697(1998);
CC -I- FUNCTION: POSSESSES A POTENT ANTIMICROBIAL ACTIVITY AGAINST
GRAM-POSITIVE AND GRAM-NEGATIVE BACTERIA. PROBABLY ACTS BY
DISTURBING MEMBRANE FUNCTIONS WITH ITS AMPHIPATHIC STRUCTURE.
CC -I- SUBCELLULAR LOCATION: SECRETED.
CC -I- SIMILARITY: BELONGS TO THE DERMASEPTIN FAMILY.
KW Antibiotic; Multigene family; Amphibian skin; Amidation.
FT MOD_RES 24 24
SQ SEQUENCE 24 AA; 2665 MW; E5987D7F50E08F4F CRC64;

Query Match 16.3%; Score 23; DB 1; Length 24;
Best Local Similarity 44.4%; Pred. NO. 2.4e+03;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 16 KDIVRSIAK 24
| | | | | | |
Db 4 KDILKNACK 12

Search completed: March 4, 2002, 13:20:28
Job time: 615 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:21:51 ; Search time 76.45 Seconds
(without alignments)
47.833 Million cell updates/sec

Title: US-09-701-623C-6
Perfect score: 141
Sequence: 1 CGETYSRVTHPLPKDIVRSIAKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 7775

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_organism.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | ID | Description |
|------------|-------|-------------|--------------|--------------------|
| 1 | 31 | 22.0 | 17 4 Q16310 | Q16310 homo sapien |
| 2 | 29 | 20.6 | 16 8 Q9T2R0 | Q9T2R0 solanum tub |
| 3 | 29 | 20.6 | 22 11 Q9QW14 | Q9QW14 mus sp. pro |
| 4 | 29 | 20.6 | 23 2 P95839 | P95839 staphylococ |
| 5 | 27 | 19.1 | 12 5 P82619 | P82619 periplaneta |
| 6 | 27 | 19.1 | 15 4 Q9UEM3 | Q9UEM3 homo sapien |
| 7 | 27 | 19.1 | 18 4 Q13665 | Q13665 homo sapien |
| 8 | 26.5 | 18.8 | 17 6 Q9XSG1 | Q9XSG1 bos taurus |
| 9 | 26 | 18.4 | 18 4 Q9BQT0 | Q9BQT0 homo sapien |
| 10 | 26 | 18.4 | 21 4 Q9BU87 | Q9BU87 homo sapien |
| 11 | 26 | 18.4 | 21 12 Q85667 | Q85667 reovirus sp |
| 12 | 25.5 | 18.1 | 23 2 Q9ZEJ4 | Q9ZEJ4 anabaena sp |
| 13 | 25 | 17.7 | 14 3 P90342 | P90342 saccharomyc |
| 14 | 25 | 17.7 | 15 13 Q90403 | Q90403 discopyge o |
| 15 | 25 | 17.7 | 17 2 Q9L8K0 | Q9L8K0 enterococcu |
| 16 | 25 | 17.7 | 17 6 Q9TR98 | Q9TR98 canis famil |
| 17 | 25 | 17.7 | 19 4 Q16271 | Q16271 homo sapien |
| 18 | 25 | 17.7 | 19 13 Q42416 | Q42416 gallus gall |
| 19 | 25 | 17.7 | 20 5 Q9U8N5 | Q9U8N5 scaptomyza |

20 25 17.7 20 5 Q9U8N2
21 25 17.7 21 4 Q9UC16
22 25 17.7 22 10 Q22501
23 25 17.7 24 4 Q16061
24 25 17.7 25 6 Q9TTG0
25 24 17.0 13 13 P82386
26 24 17.0 13 13 P82387
27 24 17.0 16 4 Q9UR83
28 24 17.0 17 2 Q34216
29 24 17.0 17 4 Q9UCS0
30 24 17.0 17 11 Q62547
31 24 17.0 17 13 P82394
32 24 17.0 17 13 P82395
33 24 17.0 17 13 P82396
34 24 17.0 18 6 Q9GJW3
35 24 17.0 18 6 Q9GJW2
36 24 17.0 18 6 Q9GJW1
37 24 17.0 19 2 P74875
38 24 17.0 20 2 Q9R4J6
39 24 17.0 20 4 Q9UCR9
40 24 17.0 20 5 Q46158
41 24 17.0 20 13 Q9PRS2
42 24 17.0 20 13 Q9PSH5
43 24 17.0 22 2 Q53519
44 24 17.0 22 2 Q85535
45 24 17.0 22 2 Q85537

ALIGNMENTS

RESULT 1

Q16310

ID Q16310

AC Q16310

DT 01-NOV-1996

DT 01-JAN-1999

DT 01-MAY-1999

DE GC*2

GN GC*2

OS Homo sapiens

OC Eukaryota

OC Mammalia

OC NCBI_TaxID=9606

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95242701

RA Kofler A., Braun A., Jenkins T., Serjeantson S.W., Cleve H.;

RT "Characterization of mutants of the vitamin-D-binding protein/group

RL South African blacks, and 2A9 from south Germany."

DR EMBL: S77130; AAD14250.1;

FT NON_TER

SQ SEQUENCE

17 AA; 1845 MW; BB26CAD60293722C CRC64;

Query Match 22.0%; Score 31; DB 4; Length 17;

Best Local Similarity 36.4%; Pred. No. 6e+02;

Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 15 PKDIVRSIAKC 25

Db 6 PKELAKLVNKC 16

RESULT 2

Q9T2R0

ID Q9T2R0

AC Q9T2R0

DT 01-MAY-2000

DT 01-MAY-2000

DT 01-MAR-2001

DT 01-MAR-2001

DT 01-MAR-2001

DE CYTOCHROME-C REDUCTASE 14 KDA SUBUNIT (EC 1.10.2.2) (FRAGMENT).
OS Solanum tuberosum (Potato).
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RX MEDLINE=94198758; PubMed=7764624;
RA Braun H.P., Kruit V., Schmitz U.K.;
RL Planta 193:99-106(1994).
SQ SEQUENCE 16 AA; 1946 MW; BBC625F8E4A4C8E7 CRC64;

Query Match 20.6%; Score 29; DB 8; Length 16;
Best Local Similarity 50.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 11 HPLPKDI 1E
| :||:|
Db 2 HOYLPEDL 9

RESULT 3
Q9QW14 PRELIMINARY; PRT; 22 AA.
AC Q9QW14;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)
DE PROTEIN TYROSINE KINASE JAK1 (FRAGMENT).
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93293238; PubMed=8514334;
RA Harpur A.G., Zimnicki A., Wilks A.F., Falk K., Rotzschke O.,
RA Rammensee H.G.;
RT "A prominent natural H-2 Kd ligand is derived from protein tyrosine
kinase JAK1.";
RL Immunol. Lett. 35:235-237(1993).
SQ SEQUENCE 22 AA; 2681 MW; D0110BD1FC3C084B CRC64;

Query Match 20.6%; Score 29; DB 11; Length 22;
Best Local Similarity 25.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 4 TYYSRVTHPLPKDIV 19
:|: :|| : :|
Db 7 SYFPEITHIVIKESV 22

RESULT 4
P95839 PRELIMINARY; PRT; 23 AA.
ID P95839;
AC P95839;
DT 01-MAY-1997 (TRENBLREL. 03, Created)
DT 01-MAY-1997 (TRENBLREL. 03, Last sequence update)
DT 01-NOV-1998 (TRENBLREL. 08, Last annotation update)
DE ORF56 (FRAGMENT).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COL;
RA MEDLINE=96427339; PubMed=8830703;
RA Wu S., de Lencastre H., Tomasz A.;

RT "Sigma-B, a putative operon encoding alternate sigma factor of
RT Staphylococcus aureus RNA polymerase: molecular cloning and DNA
RT sequencing.";
RL J. Bacteriol. 178:6036-6042(1996).
DR EMBL; Y09029; CAA71063.1; -.
FT NON_TER 1
SQ SEQUENCE 23 AA; 2541 MW; 7F47717B1767D34F CRC64;

Query Match 20.6%; Score 29; DB 2; Length 23;
Best Local Similarity 71.4%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGETYS 7
| :||:|
Db 11 CNEYLS 17

RESULT 5
P82619 PRELIMINARY; PRT; 12 AA.
ID P82619;
AC P82619;
DT 01-OCT-2000 (TRENBLREL. 15, Created)
DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)
DE PYROKININ-4 (PEA-PK-4) (FXPRL-AMIDE).
OS Periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blattodea; Blattidae; Periplaneta.
OX NCBI_TaxID=6978;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE-RETROCEBRAL COMPLEX.
RX MEDLINE=99212469; PubMed=10196736;
RA Predel R., Kellner R., Nachman R.J., Holman G.M., Rapus J., Gaede G.;
RT "Differential distribution of pyrokinin-isoforms in cerebral and
RT abdominal neurohemal organs of the American cockroach.";
RL Insect Biochem. Mol. Biol. 29:139-144(1999).
RN [2]
RP TISSUE SPECIFICITY.
RX MEDLINE=20189894; PubMed=10723010;
RA Predel R., Eckert M.;
RT "Tagma-specific distribution of FXPRLamides in the nervous system of
RT the American cockroach.";
RL J. Comp. Neurol. 419:352-363(2000).
CC -|- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
(MYOTROPIC ACTIVITY).
CC -|- TISSUE SPECIFICITY: CORPORA CARDIACA.
CC -|- MASS SPECTROMETRY: MW=1147.9; METHOD=MALDI.
CC -|- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
KW Neuropeptide; Amidation; Pyrokinin.
FT MOD_RES 12 12
SQ SEQUENCE 12 AA; 1449 MW; FA7A3049FF42CAA1 CRC64;

Query Match 19.1%; Score 27; DB 5; Length 12;
Best Local Similarity 66.7%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 13 HLPKDI 18
| :||:|
Db 2 HLPKDV 7

RESULT 6
Q9UEM3 PRELIMINARY; PRT; 15 AA.
ID Q9UEM3;
AC Q9UEM3;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-MAY-2000 (TRENBLREL. 13, Last annotation update)
DE AXONEMAL DYNEIN, HEAVY CHAIN (FRAGMENT).

GN DNAH3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Maiti A.K., Mattei M.G., Jorissen M., Volz A., Ziegler A.,
RA Bouvagnet P.;
RT "Chromosomal localization of human dynein heavy chain genes.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ132092; CAA10565.1; -
FT NON_TER 1
FT NON_TER 15
SQ SEQUENCE 15 AA; 1646 MW; 4EDFDA937C826170 CRC64;

Query Match 19.1%; Score 27; DB 4; Length 15;
Best Local Similarity 44.4%; Pred. No. 2e+03;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 16 KDIVRSIAK 24
||: :||
Db 6 KDLAKALAK 14

RESULT 7
ID Q13665 PRELIMINARY; PRT; 18 AA.
AC Q13665;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE CARDIAC MYOSIN BINDING PROTEIN (FRAGMENT).
GN MYBP-C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96083592; PubMed=7493025;
RA Watkins H., Conner D., Thierfelder L., Jarcho J.A., MacRae C.,
RA McKenna W.J., Maron B.J., Seidman J.G., Seidman C.E.;
RT "Mutations in the cardiac myosin binding protein-C gene on chromosome
RT 11 cause familial hypertrophic cardiomyopathy.";
RL Nat. Genet. 11:434-437(1995).
DR EMBL; S80805; AAB35661.1; -
FT NON_TER 1
FT NON_TER 18
SQ SEQUENCE 18 AA; 1891 MW; 3EC842CE85A19C97 CRC64;

Query Match 19.1%; Score 27; DB 4; Length 18;
Best Local Similarity 50.0%; Pred. No. 2.4e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYYSRVT 10
||: ||
Db 9 CGGIYVCRAT 18

RESULT 8
ID Q9XSG1 PRELIMINARY; PRT; 17 AA.
AC Q9XSG1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE SMCX (FRAGMENT).
GN SMCX.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Poloumienko A., Blecher S.;
RT "Exon-intron structure of SMCX and SMCY genes in bovine and swine.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF135448; AAD34440.1; -
FT NON_TER 1
FT NON_TER 17
SQ SEQUENCE 17 AA; 1927 MW; 10351B0D516D16F0 CRC64;

Query Match 18.8%; Score 26.5; DB 6; Length 17;
Best Local Similarity 42.9%; Pred. No. 2.7e+03;
Matches 6; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 12 PHLPKDIVRSIAKC 25
|: ||: |
Db 2 PEIPKGVWR-CPKC 14

RESULT 9
ID Q9BQT0 PRELIMINARY; PRT; 18 AA.
AC Q9BQT0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 2.2 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21195339; PubMed=11297743;
RA Holzmann K., Ambrosch I., Elbling L., Micksche M., Berger W.;
RT "A small upstream open reading frame causes inhibition of human major
RT vault protein expression from a ubiquitous mRNA splice variant.";
RL FEBS Lett. 494:99-104(2001).
DR EMBL; AJ291367; CAC35315.1; -
KW Hypothetical protein.
SQ SEQUENCE 18 AA; 2179 MW; 5D06F9A3F11CB828 CRC64;

Query Match 18.4%; Score 26; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.4e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PHLP 15
||||
Db 11 PHLP 14

RESULT 10
ID Q9BU87 PRELIMINARY; PRT; 21 AA.
AC Q9BU87;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:3518).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ENDOMETRIAL ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC002821; AAH02821.1; -
SQ SEQUENCE 21 AA; 2149 MW; EF66757E7B79C6EC CRC64;

Query Match 18.4%; Score 26; DB 4; Length 21;
Best Local Similarity 40.0%; Pred. No. 4e+03;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 7 SRVTHPLPK 16
| : | : | :
Db 7 SKTRPEAPR 16

RESULT 11
Q85667 PRELIMINARY; PRT; 21 AA.
AC Q85667;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE OUTER CAPSID SHELL PROTEIN SIGMA-1 (FRAGMENT).
OS Reovirus sp.
OC Viruses; dsRNA viruses; Reoviridae; unclassified Reoviridae.
OX NCBI_TaxID=1089;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=81016754; PubMed=6158163;
RA Li J.K.K., Keene J.D., Scheible P.P., Joklik W.K.;
RT "Nature of the 3'-terminal sequences of the plus and minus strands of
the SI gene of reovirus serotypes 1, 2 and 3.";
RL Virology 105:41-51(1980).
[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=81016752; PubMed=7414954;
RA Li J.K.K., Scheible P.P., Keene J.D., Joklik W.K.;
RT "The plus strand of reovirus gene S2 is identical with its in vitro
transcript.";
RL Virology 105:283-286(1980).
DR EMBL; J02325; AAA47262.1; -
FT NON_TER 21
SQ SEQUENCE 21 AA; 2398 MW; ADE58797A20D5986 CRC64;

Query Match 18.4%; Score 26; DB 12; Length 21;
Best Local Similarity 45.5%; Pred. No. 4e+03;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 12 PHLPKDIVRSI 22
| : | : | : | :
Db 3 PRLREEVRLI 13

RESULT 12
Q92EJ4 PRELIMINARY; PRT; 23 AA.
AC Q92EJ4;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE DNAJ2 PROTEIN (FRAGMENT).
GN DNAJ2.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=PCC 7120;
RA Pohl B.;
RL Thesis (1999), University of Bonn, Botanical Institute.
DR EMBL; AJ132709; CAA10746.1; -
FT NON_TER 23
SQ SEQUENCE 23 AA; 2493 MW; ACS084286BC591ED CRC64;

Query Match 18.1%; Score 25.5; DB 2; Length 23;
Best Local Similarity 60.0%; Pred. No. 5.2e+03;
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Qy 4 TYYSRV-THP 12
| : | : | : | :
Db 8 TYYSLLGLRP 17

RESULT 13
P90342 PRELIMINARY; PRT; 14 AA.
AC P90342;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE ORF YBR090C (FRAGMENT).
GN YBR090C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
RN SEQUENCE FROM N.A.
RP Feldmann H., Mannhaupt G., Schwarzlose C., Vetter I.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP MIPS;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=95112788; PubMed=7813418;
RA Feldmann H., Aigle M., Aljinovic G., Andre B., Baclet M.C., Barthe C.,
Baur A., Becam A.M., Biteau N., Boles E., Brandt T., Brendel M.,
Brueckner M., Bussereau F., Christiansen C., Contreras R., Crouzet M.,
Cziepluch C., Demolis N., Delaveau T., Doignon F., Domdey H.,
Duesterhus S., Dubois E., Dujon B., El Bakoury M., Entian K.D.,
Fuernmann M., Fiers W., Fobo G.M., Fritz C., Gassenhuber H.,
Glansdorff N., Goffeau A., Grivell L.A., De haan M., Hein C.,
Herbert C.J., Hollenberg C.P., Holmstrom K., Jacq C., Jacquet M.,
Jauniaux J.C., Jonniaux J.L., Kallesoe T., Kiesau P., Kirchraeth L.,
Koetter P., Korol S., Liebl S., Logghe M., Lohan A.J.E., Louis E.J.,
Li Z.Y., Maat M.J., Mallet L., Mannhaupt G., Messenguy F., Miosga T.,
Molmans F., Mueller S., Nasr F., Obermaier B., Perea J., Pierard A.,
Piravandi E., Pohl F.M., Pohl T.M., Potier S., Proft M., Purnelle B.,
Ramezani Rad M., Rieger M., Rose M., Schaeff-Gerstenschlaeger I.,
Scherens B., Schwarzlose C., Skala J., Slonimski P.P., Smits P.H.M.,
Souciet J.L., Steensma H.Y., Stucka R., Urrestarazu A.,
Van der Aart Q.J., Van Dyck L., Vassarotti A., Vetter I.,
Vierendeels F., Visser S., Wagner G., de Wergifosse P., Wolfe K.H.,
Zagulska M., Zimmermann F.K., Mewes H.W., Kleine K.;
RT "Complete DNA sequence of yeast chromosome II.";
RL EMBL; J. 13:5795-5809(1994).
DR EMBL; Z35957; CAA85041.1; -
DR SGD; S0000294; YBR090C.
ET NON_TER 1
SQ SEQUENCE 14 AA; 1657 MW; C6B3A4A2E8485212 CRC64;

Query Match 17.7%; Score 25; DB 3; Length 14;
Best Local Similarity 83.3%; Pred. No. 3.7e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 KDIVRS 21
| : | : | : | :
Db 4 RDIVRS 9

RESULT 14
Q90403 PRELIMINARY; PRT; 15 AA.
ID Q90403

AC Q90403;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-NOV-1998 (Tremblrel. 08, Last annotation update)
 DE AGRIN (FRAGMENT).
 OS Discopyge ommata (Electric ray).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Squala; Hypnosqualea; Pristiogaster; Batoida;
 OC Torpediniformes; Narcinoidei; Narcinidae; Discopyge.
 OX NCBI_TaxID=7785;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE-ELECTRIC LOBE;
 RX MEDLINE=95164584; PubMed=7860635;
 RA Gesemann M., Denzer A.J., Ruegg M.A.;
 RT "Acetylcholine receptor-aggregating activity of agrin isoforms and
 RT mapping of the active site";
 RL J. Cell Biol. 128:625-636(1995).
 DR EMBL; U16146; AAA64486.1; -;
 FT NON_TER 1
 FT NON_TER 15
 FT NON_TER 15
 SQ SEQUENCE 15 AA; 1680 MW; 0B881C342FBA1D9A CRC64;

Query Match 17.7%; Score 25; DB 13; Length 15;
 Best Local Similarity 50.0%; Pred. No. 4e+03;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 VTHPLPKDI 18
 || || ||
 Db 1 VTRSHLANEI 10

RESULT 15
 Q9L8K0
 ID Q9L8K0 PRELIMINARY; PRT; 17 AA.
 AC Q9L8K0;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE VANY (FRAGMENT).
 GN VANYB.
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
 OC Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=BM4281;
 RC MEDLINE=20307504; PubMed=10846225;
 RA Dahl K.H., Lundblad E.W., Roekenes T.P., Olsvik O., Sundsfjord A.;
 RT "Genetic linkage of the vanB2 gene cluster to Tn5382 in vancomycin
 RT resistant enterococci and characterization of two novel insertion
 RT sequences"; 146:1469-1479(2000).
 RL Microbiology 146:1469-1479(2000).
 DR EMBL; AF201896; AAF73375.1; -;
 FT NON_TER 17
 FT NON_TER 17
 SQ SEQUENCE 17 AA; 2149 MW; B0204025105163A0 CRC64;

Query Match 17.7%; Score 25; DB 2; Length 17;
 Best Local Similarity 57.1%; Pred. No. 4.5e+03;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 YYSRVTH 11
 || || ||
 Db 6 YHSNVNH 12

Search completed: March 4, 2002, 13:21:52
 Job time: 653 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:10:52 ; Search time 37.41 Seconds
(without alignments)
15.038 Million cell updates/sec

Title: US-09-701-623c-6
Perfect score: 141
Sequence: 1 CGETYYSRVTHPLPKDIVRSIAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 123821

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents,AA:*
1: /cgn2.6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2.6/ptodata/2/iaa/5b_COMB.pep.*
3: /cgn2.6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2.6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2.6/ptodata/2/iaa/PTCUS_COMB.pep.*
6: /cgn2.6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 108 | 76.6 | 25 | 3 | US-09-100-414B-95 |
| 2 | 108 | 76.6 | 25 | 4 | US-09-303-323-95 |
| 3 | 72 | 51.1 | 22 | 2 | US-08-232-539B-19 |
| 4 | 72 | 51.1 | 24 | 2 | US-08-232-539D-20 |
| 5 | 40 | 28.4 | 17 | 1 | US-08-218-025A-186 |
| 6 | 36 | 25.5 | 18 | 2 | US-09-017-205-52 |
| 7 | 34.5 | 24.5 | 16 | 1 | US-08-077-797A-14 |
| 8 | 34.5 | 24.5 | 16 | 5 | PCT-US94-01238-14 |
| 9 | 34 | 24.1 | 18 | 4 | US-09-177-249-298 |
| 10 | 32.5 | 23.0 | 15 | 2 | US-08-071-094B-8 |
| 11 | 32.5 | 23.0 | 17 | 1 | US-08-311-307B-9 |
| 12 | 32 | 22.7 | 17 | 2 | US-08-847-696A-9 |
| 13 | 31.5 | 22.3 | 24 | 3 | US-08-637-759B-313 |
| 14 | 31.5 | 22.3 | 24 | 3 | US-08-871-355A-313 |
| 15 | 31 | 22.0 | 24 | 3 | US-08-592-500-27 |
| 16 | 31 | 22.0 | 24 | 3 | US-08-195-006-27 |
| 17 | 31 | 22.0 | 24 | 5 | PCT-US94-07644A-27 |
| 18 | 30.5 | 21.6 | 21 | 1 | US-07-746-705A-11 |
| 19 | 30.5 | 21.6 | 21 | 2 | US-08-380-182-10 |
| 20 | 30 | 21.3 | 13 | 1 | US-08-466-285-6 |
| 21 | 30 | 21.3 | 20 | 4 | US-09-446-504-9 |
| 22 | 30 | 21.3 | 21 | 2 | US-08-480-190-80 |
| 23 | 30 | 21.3 | 21 | 2 | US-08-488-379-80 |
| 24 | 30 | 21.3 | 21 | 5 | PCT-US93-07545-80 |
| 25 | 30 | 21.3 | 23 | 2 | US-08-244-951A-3 |
| 26 | 30 | 21.3 | 23 | 4 | US-09-029-348-6 |
| 27 | 30 | 21.3 | 23 | 3 | US-08-022-324-14 |

| | | | | | | |
|----|------|------|----|---|--------------------|-------------------|
| 28 | 29 | 20.6 | 10 | 4 | US-09-385-442-18 | Sequence 18, Appl |
| 29 | 29 | 20.6 | 16 | 1 | US-08-260-582-41 | Sequence 41, Appl |
| 30 | 29 | 20.6 | 16 | 4 | US-08-602-999A-226 | Sequence 226, App |
| 31 | 29 | 20.6 | 16 | 5 | PCT-US95-05471-41 | Sequence 41, Appl |
| 32 | 29 | 20.6 | 17 | 2 | US-08-982-597A-24 | Sequence 24, Appl |
| 33 | 29 | 20.6 | 17 | 3 | US-09-136-218-24 | Sequence 24, Appl |
| 34 | 29 | 20.6 | 18 | 2 | US-09-017-205-9 | Sequence 9, Appl |
| 35 | 29 | 20.6 | 18 | 4 | US-09-029-052-5 | Sequence 9, Appl |
| 36 | 29 | 20.6 | 18 | 4 | US-08-602-999A-371 | Sequence 371, App |
| 37 | 29 | 20.6 | 19 | 1 | US-08-116-733-12 | Sequence 12, Appl |
| 38 | 29 | 20.6 | 19 | 1 | US-08-466-615-9 | Sequence 9, Appl |
| 39 | 29 | 20.6 | 19 | 1 | US-08-466-763-9 | Sequence 9, Appl |
| 40 | 29 | 20.6 | 19 | 2 | US-08-411-142A-9 | Sequence 9, Appl |
| 41 | 29 | 20.6 | 24 | 1 | US-07-976-358-24 | Sequence 24, Appl |
| 42 | 29 | 20.6 | 25 | 1 | US-07-976-358-15 | Sequence 15, Appl |
| 43 | 29 | 20.6 | 25 | 1 | US-07-976-358-18 | Sequence 18, Appl |
| 44 | 29 | 20.6 | 25 | 1 | US-07-976-358-21 | Sequence 21, Appl |
| 45 | 28.5 | 20.2 | 23 | 3 | US-08-256-747C-59 | Sequence 59, Appl |

ALIGNMENTS

RESULT 1
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-95

Query Match 76.6%; Score 108; DB 3; Length 25;
Best Local Similarity 72.0%; Pred. No. 4.1e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
DB 1 CGETYYSRVTHPLPKDIVRSIAKC 25

RESULT 2
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-C054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IEM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303.323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100.414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 76.6%; Score 108; DB 4; Length 25;
Best Local Similarity 72.0%; Pred. No. 4.1e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPLPKDIVRSIAKC 25
||||| ||||| :||| ||
DB 1 CGETYSRVTHPLPRALMRSTTKC 25

RESULT 3
US-08-232-539D-19
; Sequence 19, Application US/08232539D
; Patent No. 5965709
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; APPLICANT: Jardiou, Paula M.
; TITLE OF INVENTION: IGE Antagonists
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-232-539D-19

Query Match 51.1%; Score 72; DB 2; Length 22;
Best Local Similarity 66.7%; Pred. No. 2.5e-05;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 GETYSRVTHPLPKDIV 19
||||| ||||| :||| ||
DB 5 GETYQCRVTHPLPRALM 22

RESULT 4
US-08-232-539D-20
; Sequence 20, Application US/08232539D
; Patent No. 5965709
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; APPLICANT: Jardiou, Paula M.
; TITLE OF INVENTION: IGE Antagonists
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-232-539D-20

Query Match 51.1%; Score 72; DB 2; Length 24;
Best Local Similarity 66.7%; Pred. No. 2.8e-05;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIV 19
Db 7 GETYQCRVTHPLPRALM 24

RESULT 5

US-08-218-025A-186
Sequence 186, Application US/08218025A
Patent No. 5556744
GENERAL INFORMATION:
APPLICANT: Weiner, David B.
APPLICANT: Ugen, Kenneth E.
APPLICANT: Williams, William V.
TITLE OF INVENTION: Methods and Compositions for Diagnosing
and Treating Certain HIV Infected Patients
NUMBER OF SEQUENCES: 197
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howson and Howson
STREET: P.O. Box 457, 321 No. 5556744ristown Road
CITY: Spring House
STATE: Pennsylvania
COUNTRY: U.S.A.
ZIP: 19477
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/218,025A
FILING DATE: 24-MAR-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/891,451
FILING DATE: 29-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Bak, Mary E.
REGISTRATION NUMBER: 31,215
REFERENCE/DOCKET NUMBER: WST33A
TELEPHONE: (215) 540-9206
TELEFAX: (215) 540-5818
INFORMATION FOR SEQ ID NO: 186:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-218-025A-186

Query Match 28.4%; Score 40; DB 1; Length 17;
Best Local Similarity 43.8%; Pred. No. 3;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGEYYSRVTHPLPK 16
Db 1 CGTYFARGPGIHSK 16

RESULT 6

US-09-017-205-52
Sequence 52, Application US/09017205
Patent No. 5965357
GENERAL INFORMATION:
APPLICANT: Marsden, Howard S.
TITLE OF INVENTION: PEPTIDE STRUCTURES AND THEIR USE IN
DIAGNOSIS OF HERPES SIMPLEX VIRUS TYPE 2
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye PC
STREET: 8th Floor, 1100 No. 5965357th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/017,205
FILING DATE: 02-FEB-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mitchard, Leonard C.
REGISTRATION NUMBER: 29,009
REFERENCE/DOCKET NUMBER: 604-436
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)816-4000
TELEFAX: (703)816-4100
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide from HSV-2 glycoprotein G
FRAGMENT TYPE: internal
US-09-017-205-52

Query Match 25.5%; Score 36; DB 2; Length 18;
Best Local Similarity 75.0%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 10 THPHLPKD 17
Db 7 THPHGPAD 14

RESULT 7

US-08-077-797A-14
Sequence 14, Application US/08077797A
Patent No. 5679548
GENERAL INFORMATION:
APPLICANT: Barbas, Carlos F.
APPLICANT: Rosenblum, Jonathan
APPLICANT: Lerner, Richard A.
TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE METAL
BINDING SITES AND COMPOSITIONS THEREOF
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
Patent Counsel
STREET: 10666 No. 5679548th Torrey Pines Road, TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/077,797A
FILING DATE: 14-JUN-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/012,566
FILING DATE: 02-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCR1276P
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: Internal
US-08-077-797A-14

Query Match 24.5%; Score 34.5; DB 1; Length 16;
Best Local Similarity 43.8%; Pred. No. 22;
Matches 7; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

Qy 2 GETTYSRVTHFHLPKD 17
| : : | : : |
Db 1 GDTHRGHLRH-HLPHD 15

RESULT 8
PCT-US94-01238-14
Sequence 14, Application PC/TUS9401238
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE METAL
TITLE OF INVENTION: BINDING SITES AND COMPOSITIONS THEREOF
NUMBER OF SEQUENCES: 65
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/01238
FILING DATE: 01-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/077,797
FILING DATE: 14-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/012,566
FILING DATE: 02-FEB-1993
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
PCT-US94-01238-14

Query Match 24.5%; Score 34.5; DB 5; Length 16;
Best Local Similarity 43.8%; Pred. No. 22;
Matches 7; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

Qy 2 GETTYSRVTHFHLPKD 17
| : : | : : |
Db 1 GDTHRGHLRH-HLPHD 15

RESULT 9
US-09-177-249-298
Sequence 298, Application US/09177249
Patent No. 6229064
GENERAL INFORMATION:
APPLICANT: Fischer, Robert L.
APPLICANT: Ohad, Nir
APPLICANT: Kiyosue, Tomohiro
APPLICANT: Yadegari, Ramin
APPLICANT: Margossian, Linda
APPLICANT: Harada, John
APPLICANT: Goldberg, Robert B.
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit
TITLE OF INVENTION: Development in Plants
FILE REFERENCE: 023070-086120US
CURRENT APPLICATION NUMBER: US/09/177,249
CURRENT FILING DATE: 1998-10-22
EARLIER APPLICATION NUMBER: US 09/071,838
EARLIER FILING DATE: 1998-05-01
NUMBER OF SEQ ID NOS: 324
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 298
LENGTH: 18
TYPE: PRT
ORGANISM: Arabidopsis sp.
US-09-177-249-298

Query Match 24.1%; Score 34; DB 4; Length 18;
Best Local Similarity 50.0%; Pred. No. 30;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 16 KDIVRSIAKC 25
| : : : : |
Db 2 KSVVRNVKQC 11

RESULT 10
US-08-671-094B-8
Sequence 8, Application US/08671094B
Patent No. 5912232
GENERAL INFORMATION:
APPLICANT: Talmadge, James E.
TITLE OF INVENTION: Anti-inflammatory Polypeptide
TITLE OF INVENTION: Antagonists of Human Interleukin-8
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dann, Dorfman, Herrell and Skillman
STREET: 1601 Market Street
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2307
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/671,094B
FILING DATE: 28-JUN-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Hagan, Patrick J.
REGISTRATION NUMBER: 27,643
REFERENCE/DOCKET NUMBER: 63086FWC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215)563-4100
TELEFAX: (215)563-4044
INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: Modified-site
LOCATION: 4..6
OTHER INFORMATION: /note= "The cysteine residues
may be substituted with aminobutyric acid, homocysteine or
diaminosuberic acid."
US-08-671-094B-8

Query Match 23.0%; Score 32.5; DB 2; Length 15;
Best Local Similarity 53.8%; Pred. No. 43;
Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 1 CGETYYSRVTHPH 13
| : | | : | | |
Db 4 CIKT-YSKPFHPH 15

RESULT 11
US-08-311-307B-9
; Sequence 9, Application US/08311307B
; Patent No. 5627156
; GENERAL INFORMATION:
; APPLICANT: Talmadge, James E.
; TITLE OF INVENTION: Polypeptide Agonist Derived From Human
; TYPE OF INVENTION: Interleukin-8
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,307B
; FILING DATE: 23-SEP-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagan, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 63085
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4..6
; OTHER INFORMATION: /note= "The cysteine residues may
; be substituted with aminobutyric acid, homocysteine or
; OTHER INFORMATION: diaminosuberic acid."

US-08-311-307B-9

Query Match 23.0%; Score 32.5; DB 1; Length 17;
Best Local Similarity 53.8%; Pred. No. 50;
Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 1 CGETYYSRVTHPH 13
| : | | : | | |
Db 6 CIKT-YSKPFHPH 17

RESULT 12
US-08-847-696A-9
; Sequence 9, Application US/08847696A
; Patent No. 5877276
; GENERAL INFORMATION:
; APPLICANT: Talmadge, James E.
; TITLE OF INVENTION: Polypeptide Agonist For Human Interleukin-8
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/847,696A
; FILING DATE: 28-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/311,307
; FILING DATE: 23-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rigaut, Kathleen D.
; REGISTRATION NUMBER: P 43,047
; REFERENCE/DOCKET NUMBER: 63085C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4..6
; OTHER INFORMATION: /note= "The xaa in positions 4 and 6
; in the peptide may be aminobutyric acid, homocysteine, cys
; OTHER INFORMATION: diaminosuberic acid."
US-08-847-696A-9

Query Match 22.7%; Score 32; DB 2; Length 17;
Best Local Similarity 62.5%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 YSRVTHPH 13
| : | | |
Db 10 YSKPFHPH 17

RESULT 13
US-08-637-759B-313
; Sequence 313, Application US/08637759B
; Patent No. 5876931
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/637.759B
; FILING DATE: 03-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8795
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 313:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-637-759B-313

Query Match 22.3%; Score 31.5; DB 2; Length 24;
Best Local Similarity 50.0%; Pred. No. 1.le+02;
Matches 7; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 4 TYYSRVTHPLPKD 17
||| :|: ||| :
Db 6 TYLSDITN-HLP AE 18

RESULT 14
US-08-871-355A-313
; Sequence 313, Application US/08871355A
; Patent No. 6015669
; GENERAL INFORMATION:
; APPLICANT: David William Holden
; TITLE OF INVENTION: Identification of Genes
; NUMBER OF SEQUENCES: 501
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871.355A
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB95/02875
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: RPMS 101 CON
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8795
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEQ ID NO: 313:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-871-355A-313

Query Match 22.3%; Score 31.5; DB 3; Length 24;
Best Local Similarity 50.0%; Pred. No. 1.le+02;
Matches 7; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

Qy 4 TYYSRVTHPLPKD 17
||| :|: ||| :
Db 6 TYLSDITN-HLP AE 18

RESULT 15
US-08-592-500-27
; Sequence 27, Application US/08592500
; Patent No. 6005089
; GENERAL INFORMATION:
; APPLICANT: Lanza, Francois
; APPLICANT: Phillips, David R.
; APPLICANT: Cazenave, Jean-Pierre
; TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourile and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592.500
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,455
; FILING DATE: 09-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 12418-28
; TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Region
; LOCATION: 1..24
; OTHER INFORMATION: /note= "Tandem Leu-rich repeated
; OTHER INFORMATION: structure for platelet GPV."
; US-08-592-500-27

```

```

Query Match 22.0%; Score 31; DB 3; Length 24;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 13 HLPKDIVRSIAK 24
Db 4 HLPKGLLGAQAK 15

```

Search completed: March 4, 2002, 13:10:52
Job time: 299 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:11:39 ; Search time 40.99 Seconds
(without alignments)
46.459 Million cell updates/sec

Title: US-09-701-623c-7
Perfect score: 145
Sequence: 1 CGEGYQSRVDHPHPKPIVRSITKC 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 4959

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 33 | 22.8 | 23 | S48156 | alpha-amylase inhi |
| 2 | 29 | 20.0 | 18 | S26664 | microtubule-associ |
| 3 | 29 | 20.0 | 19 | B53145 | high conductance c |
| 4 | 28 | 19.3 | 13 | A05174 | tryptophyllin-l3 - |
| 5 | 28 | 19.3 | 23 | E39855 | paralytic peptide |
| 6 | 28 | 19.3 | 23 | C39855 | paralytic peptide |
| 7 | 28 | 19.3 | 23 | D39855 | paralytic peptide |
| 8 | 27 | 18.6 | 13 | S21152 | tryptophyllin-rela |
| 9 | 27 | 18.6 | 24 | S09363 | hypothetical WTCOI |
| 10 | 26.5 | 18.3 | 20 | S09025 | carboxylesterase (|
| 11 | 26 | 17.9 | 18 | I78841 | thrombospondin rec |
| 12 | 26 | 17.9 | 20 | A41439 | acid ribonuclease |
| 13 | 26 | 17.9 | 22 | JN0910 | M4-(beta-N-acetylq |
| 14 | 26 | 17.9 | 22 | A47415 | mannose-1-phosphat |
| 15 | 26 | 17.9 | 24 | S29749 | serum albumin - do |
| 16 | 25.5 | 17.6 | 25 | S03456 | T-cell receptor al |
| 17 | 25 | 17.2 | 13 | B47415 | mannose-1-phosphat |
| 18 | 25 | 17.2 | 16 | I51879 | cystathionine beta |
| 19 | 25 | 17.2 | 19 | T02624 | hypothetical prote |
| 20 | 25 | 17.2 | 20 | PN0133 | pepsin (EC 3.4.23. |
| 21 | 25 | 17.2 | 20 | A61093 | glue protein - Cal |
| 22 | 25 | 17.2 | 21 | I54351 | gene HEXA protein |
| 23 | 25 | 17.2 | 21 | S41390 | p7 protein - human |
| 24 | 25 | 17.2 | 23 | S72535 | probable acr-2 reg |
| 25 | 25 | 17.2 | 25 | I40592 | cena protein (IgAl |
| 26 | 25 | 17.2 | 25 | A60704 | serine proteinase |
| 27 | 24 | 16.6 | 14 | PH1597 | Ig H chain V-D-J r |
| 28 | 24 | 16.6 | 16 | T37075 | hypothetical prote |
| 29 | 24 | 16.6 | 18 | B49254 | tCr C gamma 1 chai |

30 24 16.6 18 2 B32473 histidine-rich pro
31 24 16.6 18 2 I46653 T-cell receptor de
32 24 16.6 20 2 S58382 hypothetical prote
33 24 16.6 21 2 PQ0257 microbial serine p
34 24 16.6 21 2 I50535 calmodulin - elect
35 24 16.6 22 2 I50533 calmodulin - elect
36 24 16.6 22 2 S73389 hypothetical prote
37 24 16.6 23 2 A60423 monophenol monooxy
38 24 16.6 23 2 A32473 histidine-rich pro
39 24 16.6 25 2 A18864 enkephalin-contain
40 23 15.9 8 2 B45800 serum albumin - do
41 23 15.9 10 2 H37196 bradykinin-potent
42 23 15.9 14 1 NTKN1M alpha-conotoxin MI
43 23 15.9 14 2 C35141 T-cell receptor de
44 23 15.9 15 2 PH1788 T-cell receptor al
45 23 15.9 15 2 PX0031 mixed lymphocyte r

ALIGNMENTS

RESULT 1

S48156
alpha-amylase inhibitor - rye
C:Species: Secale cereale (rye)
C:Date: 07-May-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C:Accession: S48156
R:Garcia-Casado, G.; Sanchez-Monge, R.; Lopez-Otin, C.; Salcedo, G.
Eur. J. Biochem. 224, 525-531, 1994
A:Title: Rye inhibitors of animal alpha-amylases show different specificities, aggregat
A:Reference number: S48156; MUID:95010030
A:Accession: S48156
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-23 <GAR>
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: alpha-amylase inhibitor

Query Match 22.8% Score 33; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 1.6e+02;
Matches 7; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 2 GEGYQSRVDHPHPKPIVRSI 22

Db 3 GENCWPCMGHPMPFPCRAL 23

RESULT 2

S26664
microtubule-associated protein tau - human
C:Species: Homo sapiens (man)
C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 26-May-1995
C:Accession: S26664
R:Andreadis, A.; Brown, W.M.; Kosik, K.S.
Biochemistry 31, 10626-10633, 1992
A:Title: Structure and novel exons of the human tau gene.
A:Reference number: S26662; MUID:93041757
A:Accession: S26664
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-18 <AND>
A:Cross-references: EMBL:X61375
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1991

Query Match 20.0% Score 29; DB 2; Length 18;

Best Local Similarity 71.4%; Pred. No. 4.8e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 6 QSRVDHP 12

Db 4 KSREDHP 10

RESULT 3
B53145
high conductance calcium-activated potassium channel, maxi-K channel - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C:Accession: B53145
J:Knaus, H.G.; Garcia-Calvo, M.; Kaczorowski, G.J.; Garcia, M.L.
R:Biol. Chem. 269, 3921-3924, 1994
A:Title: Subunit composition of the high conductance calcium-activated potassium channel
A:Reference number: A53145; MUID:94140798
A:Accession: B53145
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-19 <KNA>
A>Note: sequence extracted from NCBI backbone (NCBIP:144547)

Query Match 20.0%; Score 29; DB 2; Length 19;
Best Local Similarity 54.5%; Pred. No. 5.1e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 3 EGYQSRVDHPH 13
| | | | |
Db 6 EFGYGSVLNPH 16

RESULT 4
A05174
tryptophyllin-13 - Rohde's leaf frog
C:Species: Phyllomedusa rohdei (Rohde's leaf frog)
C>Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 18-Aug-2000
C:Accession: A05174
R:Montecucchi, P.C.; Gozzini, L.; Erspamer, V.
Int. J. Pept. Protein Res. 27, 175-182, 1986
A:Reference number: A05174
A:Accession: A05174
A:Molecule type: protein
A:Residues: 1-13 <MON>
C:Superfamily: unassigned animal peptides
C:Keywords: pyroglutamic acid; skin
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 19.3%; Score 28; DB 2; Length 13;
Best Local Similarity 44.4%; Pred. No. 4.8e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 10 DHPHPRPPI 18
| | | | |
Db 2 EKPYWRPPI 10

RESULT 5
E39855
paralytic peptide III - beet armyworm
C:Species: Spodoptera exigua (beet armyworm)
C>Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 30-Sep-1993
C:Accession: E39855
R:Skinner, W.S.; Dennis, P.A.; Li, J.P.; Summerfelt, R.M.; Carney, R.L.; Quistad, G.B.
J. Biol. Chem. 266, 12873-12877, 1991
A:Title: Isolation and identification of paralytic peptides from hemolymph of the lepid
A:Reference number: A39855; MUID:91302298
A:Accession: E39855
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-23 <SKI>
C:Superfamily: paralytic peptide I

Query Match 19.3%; Score 28; DB 2; Length 23;
Best Local Similarity 50.0%; Pred. No. 8.8e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGEGYQSRVD 10
| | | | |
Db 7 CTPGYQRTAD 16

RESULT 6
C39855
paralytic peptide I - beet armyworm
C:Species: Spodoptera exigua (beet armyworm)
C>Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 30-Sep-1993
C:Accession: C39855
R:Skinner, W.S.; Dennis, P.A.; Li, J.P.; Summerfelt, R.M.; Carney, R.L.; Quistad, G.B.
J. Biol. Chem. 266, 12873-12877, 1991
A:Title: Isolation and identification of paralytic peptides from hemolymph of the lepid
A:Reference number: A39855; MUID:91302298
A:Accession: C39855
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-23 <SKI>
C:Superfamily: paralytic peptide I

Query Match 19.3%; Score 28; DB 2; Length 23;
Best Local Similarity 50.0%; Pred. No. 8.8e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGEGYQSRVD 10
| | | | |
Db 7 CTPGYQRTAD 16

RESULT 7
D39855
paralytic peptide II - beet armyworm
C:Species: Spodoptera exigua (beet armyworm)
C>Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 30-Sep-1993
C:Accession: D39855
R:Skinner, W.S.; Dennis, P.A.; Li, J.P.; Summerfelt, R.M.; Carney, R.L.; Quistad, G.B.
J. Biol. Chem. 266, 12873-12877, 1991
A:Title: Isolation and identification of paralytic peptides from hemolymph of the lepid
A:Reference number: A39855; MUID:91302298
A:Accession: D39855
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-23 <SKI>
C:Superfamily: paralytic peptide I

Query Match 19.3%; Score 28; DB 2; Length 23;
Best Local Similarity 50.0%; Pred. No. 8.8e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGEGYQSRVD 10
| | | | |
Db 7 CTPGYQRTAD 16

RESULT 8
S21152
tryptophyllin-related peptide - two-colored leaf frog
C:Species: Phyllomedusa bicolor (two-colored leaf frog)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 18-Aug-2000
C:Accession: S21152
R:Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Kreil,
FEBS Lett. 302, 151-154, 1992
A:Title: Identification and characterization of two dermorphins from skin extracts of
A:Reference number: S21152; MUID:92339502
A:Accession: S21152
A:Molecule type: protein
A:Residues: 1-13 <MIG>
A:Experimental source: skin

C;Superfamily: unassigned animal peptides

Query Match 18.6%; Score 27; DB 2; Length 13;
Best Local Similarity 44.4%; Pred. No. 6.7e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 10 DHPFPKPPI 18
|||
Db 2 EKFPYPPPI 10

RESULT 9

S09363
Hypothetical MTCOL/MTCYB mutant fusion protein - human mitochondrion (fragment)

C;Species: Mitochondrion Homo sapiens (man)
C;Date: 21-Nov-1993 #sequence_revision 14-Aug-1997 #text_change 20-Apr-2000
C;Accession: S09363
R;Poulton, J.; Deadman, M.E.; Gardiner, R.M.
Nucleic Acids Res. 17, 10223-10229, 1989
A;Title: Tandem direct duplications of mitochondrial DNA in mitochondrial myopathy: anal
A;Reference number: S09363; MUID:90098864
A;Accession: S09363
A;Molecule type: DNA
A;Residues: 1-24 <POU>
C;Comment: This is the hypothetical translation of a sequence believed to result from a

C;Genetics:
A;Gene: MTCOL/MTCYB
A;Genome: Mitochondrion
A;Genetic code: SGC1
C;Keywords: fusion protein; mitochondrion
F;1-4/Region: cytochrome-c oxidase chain I
F;5-24/Region: cytochrome b (+2 frame shifted)

Query Match 18.6%; Score 27; DB 4; Length 24;
Best Local Similarity 31.2%; Pred. No. 1.3e+03;
Matches 5; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 10 DHPFPKPPIVRSITKC 25
|||
Db 6 DHFSTQKPETSALSSC 21

RESULT 10

S09025
carboxylesterase (EC 3.1.1.1), microsomal - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Aug-1998
C;Accession: S09025
R;Hosokawa, M.; Maki, T.; Satoh, T.
Arch. Biochem. Biophys. 277, 219-227, 1990
A;Title: Characterization of molecular species of liver microsomal carboxylesterases of
A;Reference number: S09021; MUID:90179180
A;Accession: S09025
A;Molecule type: protein
A;Residues: 1-20 <HOS>
C;Superfamily: cholinesterase; cholinesterase homology
C;Keywords: carboxylic ester hydrolase

Query Match 18.3%; Score 26.5; DB 2; Length 20;
Best Local Similarity 41.7%; Pred. No. 1.3e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 11 HPFPKPPIVRSI 22
|||
Db 1 HPSXP-PVNVXV 11

RESULT 11

I78841
Thrombopoietin receptor - mouse (fragment)

C;Species: Mus sp. (mouse)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C;Accession: I78841
R;Alexander, W.S.; Dunn, A.R.
Oncogene 10, 795-803, 1995

A;Title: Structure and transcription of the genomic locus encoding murine c-Mpl, a re
A;Reference number: I58350; MUID:95166571
A;Accession: I78841

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-18 <RES>

A;Cross-references: GB:S76842; NID:g912990; PIDN:AAB33462.1; PID:g912991

C;Genetics:

A;Gene: c-mpl1

Query Match 17.9%; Score 26; DB 2; Length 18;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 12 PHFPKPPIV 19
|||
Db 2 PHGPAPFL 9

RESULT 12

A41439
acid ribonuclease (EC 3.1.1.-) - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 21-Mar-1996
C;Accession: A41439
R;Ohgi, K.; Sanda, A.; Takizawa, Y.; Irie, M.
J. Biochem. 103, 267-273, 1988
A;Title: Purification of acid ribonucleases from bovine spleen.
A;Reference number: A41439; MUID:88227899
A;Accession: A41439
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <OHG>
C;Keywords: hydrolase

Query Match 17.9%; Score 26; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 11 HPFPKPPI 18
|||
Db 9 HLYFPKDL 16

RESULT 13

JN0910
N4-(beta-N-acetylglucosaminyl)-L-asparaginase (EC 3.5.1.26) alpha chain - Flavobacter
N;Alternate names: glycosylasparaginase; N4-(N-acetyl-beta-glucosaminyl)-L-asparagin
C;Species: Flavobacterium meningosepticum
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 17-Mar-1999
C;Accession: JN0910
R;Tarentino, A.L.; Plummer Jr., T.H.
Biochem. Biophys. Res. Commun. 197, 179-186, 1993
A;Title: The first demonstration of a procaryotic glycosylasparaginase.
A;Reference number: JN0910; MUID:94071939
A;Accession: JN0910
A;Molecule type: protein
A;Residues: 1-22 <TAR>
C;Comment: This heterodimeric enzyme is the counterpart to a lysosomal amidase/amidoh
gine-linked glycans.
C;Keywords: heterodimer; hydrolase

Query Match 17.9%; Score 26; DB 2; Length 22;
Best Local Similarity 75.0%; Pred. No. 1.7e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 16 KPVRISIT 23
|||||
Db 3 KPVLSTT 10

RESULT 14

A47415
mannose-1-phosphate guanylyltransferase (EC 2.7.7.13) 37K beta chain - pig (fragment)
N:Alternate names: GDP-mannose pyrophosphorylase 37K beta chain
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 25-Feb-1994 #sequence_revision 12-Aug-1996 #text_change 13-Sep-1998
C:Accession: A47415
R:Sumilo, T.; Drake, R.R.; York, J.L.; Elbein, A.D.
J. Biol. Chem. 268, 17943-17950, 1993
A:Title: GDP-mannose pyrophosphorylase. Purification to homogeneity, properties, and uti
A:Reference number: A47415; MUID:93352609
A:Contents: liver
A:Accession: A47415
A:Molecule type: protein
A:Residues: 1-23 <SZU>
A:Note: sequence extracted from NCBI backbone (NCBIP:136438)
C:Complex: The enzyme appears to be a heterodimer of alpha and beta chains.
C:Function:
A:Description: generates GDP-mannose and pyrophosphate from mannose-1-phosphate and GTP
A:Note: also catalyzes synthesis of GDP-glucose from glucose-1-phosphate (EC 2.7.7.34 ac
C:Superfamily: mannose-1-phosphate guanylyltransferase
C:Keywords: nucleotidyltransferase

Query Match 17.9%; Score 26; DB 2; Length 23;
Best Local Similarity 33.3%; Pred. No. 1.7e+03;
Matches 7; Conservative 4; Mismatches 4; Indels 6; Gaps 1;

Qy 4 GYQSRVDHPHPKPVRISITK 24
|||:
Db 9 GYGTSL-----RPLTSLSPK 23

RESULT 15

S29749
serum albumin - dog
C:Species: Canis lupus familiaris (dog)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C:Accession: S29749
R:Dixon, J.W.; Sarkar, B.
J. Biol. Chem. 249, 5872-5877, 1974
A:Title: Isolation, amino acid sequence and copper(II)-binding properties of peptide (1
A:Reference number: S29749; MUID:75011422
A:Accession: S29749
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-24 <DIX>
C:Superfamily: serum albumin; serum albumin repeat homology

Query Match 17.9%; Score 26; DB 2; Length 24;
Best Local Similarity 44.4%; Pred. No. 1.8e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 EGYQSFVDH 11
|:
Db 1 EAYKSEIAH 9

Search completed: March 4, 2002, 13:11:39
Job time: 291 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:20:28 ; Search time 24.21 Seconds
(without alignments)
37.861 Million cell updates/sec

Title: US-09-701-623C-7
Perfect score: 145
Sequence: 1 CGEGYQSRVDHPHFKPKIVRSITKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1446

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 28 | 19.3 | 13 | 1 | TP13_PHYRO |
| 2 | 28 | 19.3 | 23 | 1 | CP23_SPOER |
| 3 | 28 | 19.3 | 23 | 1 | PAP1_SPOEX |
| 4 | 28 | 19.3 | 23 | 1 | PAP2_SPOEX |
| 5 | 28 | 19.3 | 23 | 1 | PAP3_SPOEX |
| 6 | 25.5 | 17.6 | 16 | 1 | MK2B_PALPR |
| 7 | 24 | 16.6 | 15 | 1 | MCA2_RHOOP |
| 8 | 23.5 | 16.2 | 15 | 1 | MK1_PALPR |
| 9 | 23.5 | 16.2 | 15 | 1 | MK2A_PALPR |
| 10 | 23.5 | 16.2 | 16 | 1 | MK3_PALPR |
| 11 | 23 | 15.9 | 9 | 1 | NEUX_HUMAN |
| 12 | 23 | 15.9 | 9 | 1 | NEUX_RAT |
| 13 | 23 | 15.9 | 10 | 1 | BPP8_BOTIN |
| 14 | 23 | 15.9 | 14 | 1 | CXAL_CONMA |
| 15 | 23 | 15.9 | 19 | 1 | HI70_RAT |
| 16 | 23 | 15.9 | 20 | 1 | PIRR_PYRAP |
| 17 | 23 | 15.9 | 23 | 1 | PAP1_MANSE |
| 18 | 22.5 | 15.5 | 21 | 1 | THAN_PODMA |
| 19 | 22.5 | 15.5 | 22 | 1 | TX12_TRIWA |
| 20 | 22.5 | 15.5 | 25 | 1 | CXAA_CONPU |
| 21 | 22 | 15.2 | 10 | 1 | BPP2_BOTIN |
| 22 | 22 | 15.2 | 19 | 1 | AMY_DERPT |
| 23 | 22 | 15.2 | 21 | 1 | CFPA_TREPH |
| 24 | 22 | 15.2 | 21 | 1 | LPRM_CORDI |
| 25 | 22 | 15.2 | 24 | 1 | BRIA_RANES |
| 26 | 22 | 15.2 | 24 | 1 | CXOA_CONST |
| 27 | 22 | 15.2 | 25 | 1 | H2BI_ECHES |
| 28 | 21.5 | 14.8 | 15 | 1 | MAOX_CHICK |
| 29 | 21 | 14.5 | 13 | 1 | MLA_ANOCA |
| 30 | 21 | 14.5 | 13 | 1 | MLA_CAMDR |
| 31 | 21 | 14.5 | 14 | 1 | H4_EUPCR |
| 32 | 21 | 14.5 | 14 | 1 | MY14_EISFO |
| 33 | 21 | 14.5 | 14 | 1 | MY14_PHEVI |

RESULT 1
TP13_PHYRO

| ID | TP13_PHYRO | STANDARD | PRT | 13 AA |
|----|-------------------------------------------------------------------|----------|-----|-------|
| AC | P04096; | | | |
| DT | 01-NOV-1986 (Rel. 03, Created) | | | |
| DT | 01-NOV-1986 (Rel. 03, Last sequence update) | | | |
| DT | 01-APR-1988 (Rel. 07, Last annotation update) | | | |
| DE | TRYPTOPHYLLIN-13. | | | |
| OS | Phyllomedusa rohdei (Rohde's leaf frog). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae; | | | |
| OC | Phyllomedusa. | | | |
| OX | NCBI_TaxID=8394; | | | |
| RN | [1] | | | |
| RP | SEQUENCE. | | | |
| RA | Montecucchi P.C., Gozzini L., Erspamer V.; | | | |
| RT | "Primary structure determination of a tryptophan-containing | | | |
| RT | tridecapeptide from Phyllomedusa rohdei." | | | |
| RL | Int. J. Pept. Protein Res. 27:175-182(1986). | | | |
| DR | PIR; A05174; A05174. | | | |
| KW | Amphibian skin. | | | |
| FT | MOD_RES | | | |
| SQ | SEQUENCE 13 AA; 1646 MW; 33BF33A212227773 CRC64; | | | |

Query Match 19.3%; Score 28; DB 1; Length 13;
Best Local Similarity 44.4%; Pred. No. 2.4e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 10 DHPHFPPKI 18
:|::|||
Db 2 EKPWPPPI 10

RESULT 2
CP23_SPOER

| ID | CP23_SPOER | STANDARD | PRT | 23 AA |
|----|----------------------------------------------------------------------|----------|-----|-------|
| AC | P56683; | | | |
| DT | 15-JUL-1999 (Rel. 38, Created) | | | |
| DT | 15-JUL-1999 (Rel. 38, Last sequence update) | | | |
| DT | 15-JUL-1999 (Rel. 38, Last annotation update) | | | |
| DE | CARDIOACTIVE PEPTIDE CAP23. | | | |
| OS | Spodoptera eridania (Southern armyworm). | | | |
| OC | Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta; | | | |
| OC | Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; | | | |
| OC | Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera. | | | |
| OX | NCBI_TaxID=37547; | | | |
| RN | [1] | | | |
| RP | SEQUENCE. | | | |
| RA | MEDLINE=99196260; PubMed=10098624; | | | |
| RA | Furuya K., Hackett M., Cirelli M.A., Schegg K.M., Wang H., | | | |
| RA | Shabanowitz J., Hunt D.F., Schooley D.A.; | | | |
| RT | "A cardioactive peptide from the southern armyworm, Spodoptera | | | |
| RT | eridania." | | | |
| RL | Peptides 20:53-61(1999). | | | |
| CC | -I- FUNCTION: HAS EXCITATORY EFFECTS ON A SEMI-ISOLATED HEART FROM | | | |

CC LARVAL MANDUCA SEXTA, CAUSING AN INOTROPIC EFFECT AT LOW
 CC CONCENTRATIONS OF PEPTIDE AND CHRONOTROPIC AND INOTROPIC EFFECTS
 CC AT HIGH DOSES.
 CC -1- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
 DR InterPro: IPR003463; GBP_PSP.
 DR Pfam: PF02425; GBP_PSP; 1.
 FT DISULFID 7 19 BY SIMILARITY.
 SQ SEQUENCE 23 AA: 2519 MW; 0A96D72A70855AE0 CRC64;

Query Match 19.3%; Score 28; DB 1; Length 23;
 Best Local Similarity 50.0%; Pred. No. 4.3e+02;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVD 10
 | | | | |
 DB 7 CTPGYQRTAD 16

RESULT 3
 PAP2_SPOEX
 ID PAP1_SPOEX STANDARD; PRT; 23 AA.
 AC P30255;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PARALYTIC PEPTIDE I (PP I).
 OS Spodoptera exigua (Beet armyworm).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 CC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
 OX NCBI_TaxID=7107;
 RN [1]

RP TISSUE=Hemolymph;
 RC MEDLINE=91302298; PubMed=2071576;
 RX Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
 RA Quistad G.B.;
 RT "Isolation and identification of paralytic peptides from hemolymph of
 RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
 RT Heliothis virescens.";
 RL J. Biol. Chem. 266:12873-12877(1991).
 CC -1- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
 CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
 CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
 CC -1- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
 CC PIR: C39855; C39855.
 DR InterPro: IPR003463; GBP_PSP.
 DR Pfam: PF02425; GBP_PSP; 1.
 KW Hemolymph.
 FT DISULFID 7 19 BY SIMILARITY.
 SQ SEQUENCE 23 AA: 2451 MW; 0A96D1F60855AE0 CRC64;

Query Match 19.3%; Score 28; DB 1; Length 23;
 Best Local Similarity 50.0%; Pred. No. 4.3e+02;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVD 10
 | | | | |
 DB 7 CTPGYQRTAD 16

RESULT 4
 PAP2_SPOEX
 ID PAP2_SPOEX STANDARD; PRT; 23 AA.
 AC P30256;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PARALYTIC PEPTIDE II (PP II).
 OS Spodoptera exigua (Beet armyworm).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
 OX NCBI_TaxID=7107;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Hemolymph;
 RX MEDLINE=91302298; PubMed=2071576;
 RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
 RA Quistad G.B.;
 RT "Isolation and identification of paralytic peptides from hemolymph of
 RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
 RT Heliothis virescens.";
 RL J. Biol. Chem. 266:12873-12877(1991).
 CC -1- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
 CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
 CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
 CC -1- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
 CC PIR: D39855; D39855.
 DR InterPro: IPR003463; GBP_PSP.
 DR Pfam: PF02425; GBP_PSP; 1.
 KW Hemolymph.
 FT DISULFID 7 19 BY SIMILARITY.
 SQ SEQUENCE 23 AA: 2477 MW; 0A96CB460855AE0 CRC64;

Query Match 19.3%; Score 28; DB 1; Length 23;
 Best Local Similarity 50.0%; Pred. No. 4.3e+02;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVD 10
 | | | | |
 DB 7 CTPGYQRTAD 16

RESULT 5
 PAP3_SPOEX
 ID PAP3_SPOEX STANDARD; PRT; 23 AA.
 AC P30257;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PARALYTIC PEPTIDE III (PP III).
 OS Spodoptera exigua (Beet armyworm).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 CC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
 OX NCBI_TaxID=7107;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Hemolymph;
 RX MEDLINE=91302298; PubMed=2071576;
 RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
 RA Quistad G.B.;
 RT "Isolation and identification of paralytic peptides from hemolymph of
 RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
 RT Heliothis virescens.";
 RL J. Biol. Chem. 266:12873-12877(1991).
 CC -1- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
 CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
 CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
 CC -1- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
 CC PIR: E39855; E39855.
 DR InterPro: IPR003463; GBP_PSP.
 DR Pfam: PF02425; GBP_PSP; 1.
 KW Hemolymph.
 FT DISULFID 7 19 BY SIMILARITY.
 SQ SEQUENCE 23 AA: 2505 MW; 0A96CB5EB7D55AE0 CRC64;

Query Match 19.3%; Score 28; DB 1; Length 23;
 Best Local Similarity 50.0%; Pred. No. 4.3e+02;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGRGYQSRVD 10
| . | | | |
Db 7 CTGCGYQRTAD 16

RESULT 6
MK2B_PALPR
ID MK2B_PALPR STANDARD; PRT; 16 AA.
AC P80410;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE METALNIKOWIN IIB.
OS Palomena prasina.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomidae;
OC Pentatomidae; Palomena.
OX NCBI_TaxID=55431;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RA Chernysh S.; Cociancich S.; Briand J.-P., Hetru C., Bulet P.;
RT "The inducible antibacterial peptides of the hemipteran insect
Palomena prasina. Identification of a unique family of proline-rich
peptides and of a novel insect defensin.";
RL J. Insect Physiol. 42:81-89(1996).
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
BACTERIA.
CC -!- INDUCTION: BY BACTERIAL INFECTION.
KW Antibiotic; Insect immunity.
SQ SEQUENCE 16 AA; 2040 MW; AF21407D063B9462 CRC64;

Query Match 17.6%; Score 25.5; DB 1; Length 16;
Best Local Similarity 40.08; Pred. No. 6.9e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

QY 9 VDHHPH-PKPVRSI 22
| | | | | | | |
Db 1 VDKPDYRPRPNN 15

RESULT 7
MCA2_RHOOP
ID MCA2_RHOOP STANDARD; PRT; 15 AA.
AC P56870;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE MALEYLACETATE REDUCTASE II (EC 1.3.1.32) (FRAGMENT).
OS Rhodococcus opacus (Nocardia opaca).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OX Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
RN NCBI_TaxID=37919;
RP SEQUENCE.
RC STRAIN=ICP;
RA MEDLINE=98324954; PubMed=9657989;
RA Seibert V., Kourbatova E.M., Golovleva L.A., Schloemann M.;
RT "Characterization of a maleylacetate reductase encoding region from
Rhodococcus opacus ICP.";
RL J. Bacteriol. 180:3503-3508(1998).
CC -!- CATALYTIC ACTIVITY: 3-OXOADIPATE + NAD(P)(+) = 2-MALEYLACETATE +
NAD(P)H.
CC -!- PATHWAY: 3-CHLOROCATECHOL DEGRADATION (BETA-KETOADIPATE PATHWAY).
CC THIS PATHWAY SERVES A VITAL ROLE IN THE BIODEGRADATION OF TOXIC
AROMATIC COMPOUNDS INTRODUCED IN THE ENVIRONMENT BOTH AS NATURAL
PRODUCTS AND AS INDUSTRIAL EFFLUENT.
CC -!- SIMILARITY: BELONGS TO THE IRON-CONTAINING ALCOHOL DEHYDROGENASE
FAMILY.
CR InterPro; IPR001670; Fe-ADH.
DR PROSITE; PS00913; ADH_IRON_1; PARTIAL.

DR PROSITE; PS00060; ADH_IRON_2; PARTIAL.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.
FT NON_TER 15
SQ SEQUENCE 15 AA; 1884 MW; 58DA90DD038F025E CRC64;

Query Match 16.6%; Score 24; DB 1; Length 15;
Best Local Similarity 36.4%; Pred. No. 1.1e+03;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 RVDHPHPKPPI 18
| : | : | : |
Db 2 RFEHENLPQRI 12

RESULT 8
MK1_PALPR
ID MK1_PALPR STANDARD; PRT; 15 AA.
AC P80408;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE METALNIKOWIN I.
OS Palomena prasina.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomidae;
OC Pentatomidae; Palomena.
OX NCBI_TaxID=55431;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RA Chernysh S.; Cociancich S.; Briand J.-P., Hetru C., Bulet P.;
RT "The inducible antibacterial peptides of the hemipteran insect
Palomena prasina. Identification of a unique family of proline-rich
peptides and of a novel insect defensin.";
RL J. Insect Physiol. 42:81-89(1996).
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
BACTERIA.
CC -!- INDUCTION: BY BACTERIAL INFECTION.
KW Antibiotic; Insect immunity.
SQ SEQUENCE 15 AA; 1838 MW; 21407E663CE46299 CRC64;

Query Match 16.2%; Score 23.5; DB 1; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.3e+03;
Matches 5; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 9 VDHHPH-PKP 17
| | | | | | |
Db 1 VDKPDYRPRP 10

RESULT 9
MK2A_PALPR
ID MK2A_PALPR STANDARD; PRT; 15 AA.
AC P80409;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE METALNIKOWIN IIA.
OS Palomena prasina.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda;
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomidae;
OC Pentatomidae; Palomena.
OX NCBI_TaxID=55431;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RA Chernysh S.; Cociancich S.; Briand J.-P., Hetru C., Bulet P.;
RT "The inducible antibacterial peptides of the hemipteran insect
Palomena prasina. Identification of a unique family of proline-rich
peptides and of a novel insect defensin.";
RL J. Insect Physiol. 42:81-89(1996).
CC -!- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
BACTERIA.
CC -!- INDUCTION: BY BACTERIAL INFECTION.
KW Antibiotic; Insect immunity.
SQ SEQUENCE 15 AA; 1838 MW; 21407E663CE46299 CRC64;

RT peptides and of a novel insect defensin.";
 RL J. Insect Physiol. 42:81-89(1996).
 CC -|- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
 CC BACTERIA.
 CC -|- INDUCTION: BY BACTERIAL INFECTION.
 KW Antibiotic; Insect immunity.
 SQ SEQUENCE 15 AA; 1893 MW; 23835D063B946299 CRC64;

Query Match 16.2%; Score 23.5; DB 1; Length 15;
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;
 Matches 5; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 9 VDHPIHF-PPK 17
 ||| : |||
 Db 1 VDKPDYRPP 10

RESULT 10
 MK3_PALPR STANDARD; PRT; 16 AA.
 AC P80411;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE METALNIKOWIN III.
 OS Palomena prasina.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
 OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomioidea;
 OC Pentatomidae; Palomena.
 OX NCBI_TaxID=55431;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Hemolymph;
 RA Chernysh S.; Cocciandich S.; Briand J.-P.; Hetru C.; Bulet P.;
 RT "The inducible antibacterial peptides of the hemipteran insect
 RT Palomena prasina. Identification of a unique family of proline-rich
 RT peptides and of a novel insect defensin."
 RL J. Insect Physiol. 42:81-89(1996).
 CC -|- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
 CC BACTERIA.
 CC -|- INDUCTION: BY BACTERIAL INFECTION.
 KW Antibiotic; Insect immunity.
 SQ SEQUENCE 16 AA; 2024 MW; A9E3835D063B9452 CRC64;

Query Match 16.2%; Score 23.5; DB 1; Length 16;
 Best Local Similarity 50.0%; Pred. No. 1.4e+03;
 Matches 5; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 9 VDHPIHF-PPK 17
 ||| : |||
 Db 1 VDKPDYRPP 10

RESULT 11
 NEUX_HUMAN STANDARD; PRT; 9 AA.
 AC P04277;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DT 01-JUL-1989 (Rel. 11, Last annotation update)
 DE NEUROTENSIN-RELATED PEPTIDE (NRP) (KINETENSIN).
 OS Homo sapiens (Human), Bos taurus (Bovine), and
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606, 9913, 9986;
 RN [1]
 RP SEQUENCE.

RC SPECIES=Human;
 RX MEDLINE=86242180; PubMed=3087352;

RA Mogard M.H., Kobayashi R., Chen C.F., Lee.
 RA Shively J.E., Walsh J.H.;
 RT "The amino acid sequence of kinetensin, a novel
 RT pepsin-treated human plasma: homology with human
 RT neurotensin and angiotensin."
 RL Biochem. Biophys. Res. Commun. 136:983-988(1986).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=Human, Bovine, and Rabbit;
 RX MEDLINE=87194805; PubMed=2437111;
 RA Caraway R.E., Mitra S.P., Cochrane D.E.;
 RT "Structure of a biologically active neurotensin-related peptide
 RT obtained from pepsin-treated albumin(s)."
 RL J. Biol. Chem. 262:5968-5973(1987).
 CC -|- FUNCTION: REGULATION OF FAT DIGESTION, LIPID ABSORPTION, AND
 CC BLOOD FLOW (POTENTIAL).
 CC -|- SIMILARITY: STRUCTURALLY WITH NEUROTENSIN AND ANGIOTENSIN I.
 CC STRONG SEQUENCE HOMOLOGY WITH RAT NRP.
 DR PIR; A03239; ABHUSK.
 DR PIR; A26693; A26693.
 KW Hormone.
 SQ SEQUENCE 9 AA; 1172 MW; C804DB4761F4140D CRC64;

Query Match 15.9%; Score 23; DB 1; Length 9;
 Best Local Similarity 75.0%; Pred. No. 1e+05;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 HPHF 14
 ||| : |||
 Db 5 HPYF 8

RESULT 12
 NEUX_RAT STANDARD; PRT; 9 AA.
 AC P11382;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 01-JUL-1989 (Rel. 11, Last annotation update)
 DE NEUROTENSIN-RELATED PEPTIDE (NRP).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Plasma;
 RX MEDLINE=87194805; PubMed=2437111;
 RA Caraway R.E., Mitra S.P., Cochrane D.E.;
 RT "Structure of a biologically active neurotensin-related peptide
 RT obtained from pepsin-treated albumin(s)."
 RL J. Biol. Chem. 262:5968-5973(1987).
 CC -|- FUNCTION: REGULATION OF FAT DIGESTION, LIPID ABSORPTION, AND
 CC BLOOD FLOW (POTENTIAL).
 CC -|- SIMILARITY: STRUCTURALLY WITH NEUROTENSIN AND ANGIOTENSIN I.
 CC STRONG SEQUENCE HOMOLOGY WITH HUMAN, BOVINE AND CANINE NRP.
 KW Hormone.
 SQ SEQUENCE 9 AA; 1158 MW; D0A7DB4761F4140D CRC64;

Query Match 15.9%; Score 23; DB 1; Length 9;
 Best Local Similarity 75.0%; Pred. No. 1e+05;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 11 HPHF 14
 ||| : |||
 Db 5 HPYF 8

RESULT 13
 BPP8_BOTIN STANDARD; PRT; 10 AA.
 ID BPP8_BOTIN

P30426;
01-APR-1993 (Rel. 25, Created)
01-FEB-1994 (Rel. 28, Last sequence update)
01-FEB-1994 (Rel. 28, Last annotation update)
DE BRADYKININ-POTENTIATING PEPTIDE S5.1 (ANGIOTENSIN-CONVERTING
ENZYME INHIBITOR).
OS Bothrops insularis (Island Jararaca) (Queimada Jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=90351557; PubMed=2386615;
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
RT "Primary structure and biological activity of bradykinin potentiating
peptides from Bothrops insularis snake venom.";
RL J. Protein Chem. 9:221-227(1990).
CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
CC PTR: H37196; H37196.
DR Hypotensive agent; Venom.
KW Hypotensive agent; Venom.
FT MOD_RES 1 1
SQ SEQUENCE 10 AA: 1173 MW: 2FF835545761F6D8 CRC64;

Query Match 15.9%; Score 23; DB 1; Length 10;
Best Local Similarity 60.0%; Pred. No. 1e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 HPHPF 15
Db 5 HPNIP 9
II: I

RESULT 14
CXAL_CONMA STANDARD; PRT; 14 AA.
AC P01521;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE ALPHA-CONOTOXIN MI (M1).
OS Conus magus (Magus cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=6492;
RN [1]
RP SEQUENCE.
RX MEDLINE=83073458; PubMed=7149738;
RA McIntosh J.M., Cruz L.J., Hunkapiller M.W., Gray W.R., Olivera B.M.;
RT "Isolation and structure of a peptide toxin from the marine snail
Conus magus.";
RL Arch. Biochem. Biophys. 218:329-334(1982).
RN [2]
RP DISULFIDE BONDS.
RX MEDLINE=84032400; PubMed=6630187;
RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;
RT "Conotoxin MI. Disulfide bonding and conformational states.";
RL J. Biol. Chem. 258:12247-12251(1983).
RN [3]
RP REVIEW.
RX MEDLINE=89024586; PubMed=3052286;
RA Gray W.R., Olivera B.M., Cruz L.J.;
RT "Peptide toxins from venomous Conus snails.";
RL Annu. Rev. Biochem. 57:665-700(1988).
CC -1- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
CC INHIBIT THEM.
DR PIR: A01784; NTKNIM.

KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
KW Venom.
FT DISULFID 3 8
FT DISULFID 4 14
FT MOD_RES 14 14
SQ SEQUENCE 14 AA: 1499 MW: DEEE91898BF5E5BD CRC64;
AMIDATION.

Query Match 15.9%; Score 23; DB 1; Length 14;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEGY 5
Db 8 CGKNY 12
II: I

RESULT 15
HI70_RAT STANDARD; PRT; 19 AA.
ID HI70_RAT
AC P21794;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DE HORMONE-INDUCED PROTEIN 70 KDA (HIP-70) (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RX MEDLINE=90208308; PubMed=2181662;
RA Mobbs C.V., Fink G., Pfaff D.W.;
RT "HIP-70: a protein induced by estrogen in the brain and LH-RH in the
RT pituitary.";
RL Science 247:1477-1479(1990).
CC -1- INDUCTION: MOST PROMINENT PROTEIN INDUCED BY ESTROGEN IN
CC HYPOTHALAMUS AND MOST PROMINENT PROTEIN INDUCED BY LH-RH IN
CC PITUITARY.
FT NON_TER 19 19
SQ SEQUENCE 19 AA: 2113 MW: F793A98720B68E3C CRC64;

Query Match 15.9%; Score 23; DB 1; Length 19;
Best Local Similarity 57.1%; Pred. No. 1.9e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 EGYQSRV 9
Db 8 ENFESRV 14
II: III

Search completed: March 4, 2002, 13:20:28
Job time: 615 sec

GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:21:52 ; Search time 76.45 Seconds
(without alignments)
47.833 Million cell updates/sec

Title: US-09-701-623c-7

Perfect score: 145

Sequence: 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 7775

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17.*

1: sp-archaea.*

2: sp-bacteria.*

3: sp-fungi.*

4: sp-human.*

5: sp-invertebrate.*

6: sp-mammal.*

7: sp-mhc.*

8: sp-organelle.*

9: sp-phage.*

10: sp-plant.*

11: sp-rodent.*

12: sp-virus.*

13: sp-vertebrate.*

14: sp-unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 32 | 22.1 | 22 | 4 Q13659 | Q13659 homo sapien |
| 2 | 30 | 20.7 | 17 | 4 Q16310 | Q16310 homo sapien |
| 3 | 30 | 20.7 | 18 | 2 Q56610 | Q56610 vibrio chol |
| 4 | 30 | 20.7 | 19 | 13 P87468 | P87468 xenopus lae |
| 5 | 30 | 20.7 | 22 | 11 Q9QV92 | Q9QV92 rattus sp. |
| 6 | 29 | 20.0 | 20 | 2 Q52918 | Q52918 rhizobium m |
| 7 | 28.5 | 19.7 | 23 | 12 Q9QSP0 | Q9QSP0 australian |
| 8 | 28 | 19.3 | 14 | 4 Q9P2A2 | Q9P2A2 homo sapien |
| 9 | 28 | 19.3 | 21 | 2 Q44041 | Q44041 anabaena fl |
| 10 | 28 | 19.3 | 23 | 12 Q85496 | Q85496 bovine leuk |
| 11 | 28 | 19.3 | 25 | 4 Q9UKK1 | Q9UKK1 homo sapien |
| 12 | 28 | 19.3 | 25 | 5 Q9BW01 | Q9BW01 hydra litto |
| 13 | 28 | 19.3 | 25 | 6 Q9TRY2 | Q9TRY2 sus sp. ins |
| 14 | 27 | 18.6 | 11 | 5 Q9NL65 | Q9NL65 ascaris suu |
| 15 | 27 | 18.6 | 15 | 4 Q9UCC0 | Q9UCC0 homo sapien |
| 16 | 27 | 18.6 | 17 | 2 Q34216 | Q34216 pseudomonas |
| 17 | 26 | 17.9 | 17 | 6 Q9TF98 | Q9TF98 canis famil |
| 18 | 26 | 17.9 | 22 | 7 Q9MX47 | Q9MX47 oryzias lat |
| 19 | 26 | 17.9 | 22 | 11 Q35905 | Q35905 mus musculu |

| | | | | | |
|----|----|------|----|-----------|--------------------|
| 20 | 26 | 17.9 | 23 | 6 Q9TRF4 | Q9TRF4 sus scrofa |
| 21 | 26 | 17.9 | 24 | 3 Q9UR88 | Q9UR88 aspergillus |
| 22 | 26 | 17.9 | 24 | 12 Q9TNPS | Q9TNPS human immun |
| 23 | 26 | 17.9 | 25 | 4 Q16092 | Q16092 homo sapien |
| 24 | 25 | 17.2 | 15 | 6 Q9TQ09 | Q9TQ09 bos taurus |
| 25 | 25 | 17.2 | 15 | 6 Q9TR40 | Q9TR40 bos taurus |
| 26 | 25 | 17.2 | 16 | 2 Q9LAP2 | Q9LAP2 enterococcu |
| 27 | 25 | 17.2 | 16 | 4 Q16350 | Q16350 homo sapien |
| 28 | 25 | 17.2 | 17 | 2 Q9X5I5 | Q9X5I5 enterococcu |
| 29 | 25 | 17.2 | 17 | 2 Q9WM21 | Q9WM21 enterococcu |
| 30 | 25 | 17.2 | 17 | 2 Q9WM20 | Q9WM20 enterococcu |
| 31 | 25 | 17.2 | 19 | 2 Q33755 | Q33755 streptococc |
| 32 | 25 | 17.2 | 19 | 10 Q80997 | Q80997 arabidopsis |
| 33 | 25 | 17.2 | 20 | 11 Q61871 | Q61871 mus musculu |
| 34 | 25 | 17.2 | 21 | 4 Q16017 | Q16017 homo sapien |
| 35 | 25 | 17.2 | 21 | 4 Q9UCC5 | Q9UCC5 homo sapien |
| 36 | 25 | 17.2 | 21 | 12 Q9YQ58 | Q9YQ58 porcine cir |
| 37 | 25 | 17.2 | 21 | 12 Q56126 | Q56126 porcine cir |
| 38 | 25 | 17.2 | 21 | 12 Q93191 | Q93191 porcine cir |
| 39 | 25 | 17.2 | 23 | 12 Q65291 | Q65291 human adeno |
| 40 | 25 | 17.2 | 24 | 2 Q9R5H2 | Q9R5H2 lactobacill |
| 41 | 25 | 17.2 | 25 | 4 Q16498 | Q16498 homo sapien |
| 42 | 24 | 16.6 | 16 | 8 Q34374 | Q34374 daphnia pul |
| 43 | 24 | 16.6 | 17 | 2 Q9R4Y9 | Q9R4Y9 alcaligenes |
| 44 | 24 | 16.6 | 17 | 2 Q9L8K0 | Q9L8K0 enterococcu |
| 45 | 24 | 16.6 | 20 | 2 Q9R4J6 | Q9R4J6 pseudomonas |

ALIGNMENTS

RESULT 1

ID Q13659 PRELIMINARY; PRT; 22 AA.
AC Q13659;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE CARDIAC MYOSIN BINDING PROTEIN-C (FRAGMENT).

GN MYBP-C.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OC NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96083593; PubMed=7493026;

RA Bonne G., Carrier L., Bercovici J., Cruaud C., Richard P., Hainque B.,

RA Gautel M., Labelt S., James M., Beckmann J.;

RT "Cardiac myosin binding protein-C gene splice acceptor site mutation

RT is associated with familial hypertrophic cardiomyopathy.";

RL Nat. Genet. 11:438-440(1995).

DR EMBL: S80782; AAB35663.1; .

FT NON_TER 1

SQ SEQUENCE 22 AA; 2563 MW; 5BA92048F76774CA CRC64;

Query Match 22.1%; Score 32; DB 4; Length 22;

Best Local Similarity 38.9%; Pred.No. 4.8e+02;

Matches 7; Conservative 1; Mismatches 10; Indels 0; Caps 0;

QY 8 RVDHPHPKPIVRSITKC 25

| | | | |

Db 2 RQDPGLPRPHRHGCG 19

RESULT 2

ID Q16310 PRELIMINARY; PRT; 17 AA.
AC Q16310;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-JAN-1999 (TREMBlrel. 09, Last sequence update)

DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)

DE GC*2 PROTEIN (FRAGMENT).
 GN GC*2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95242701; PubMed=7725672;
 RA Kofler A., Braun A., Jenkins T., Serjeantson S.W., Cleve H.:
 RT "Characterization of mutants of the vitamin-D-binding protein/group
 RT specific component: GC aborigine (1A1) from Australian aborigines and
 RT South African blacks, and 2A9 from south Germany.";
 RL Vox Sang. 68:50-54(1995).
 DR EMBL: S77130; AND14250.1; -.
 FT NON_TER 17 17
 SQ SEQUENCE 17 AA; 1845 MW; BB26CAD60293722C CRC64;

Query Match 20.7%; Score 30; DB 4; Length 17;
 Best Local Similarity 36.48; Pred. No. 7.4e+02;
 Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
 QY 15 PKPIVRSITK 25
 || : : ||
 Db 6 PKELAKLVNK 16

RESULT 3
 Q56610 PRELIMINARY; PRT; 18 AA.
 AC Q56610;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE ACQA (FRAGMENT).
 GN ACQA.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C6706;
 RX MEDLINE=97074686; PubMed=8917113;
 RA Franco A., Peir-En Y., Johnson J., Barry E.M., Guerra H., Maurer R.,
 RA Morris J.G.:
 RT "Cloning and characterization of dnaE, encoding the catalytic subunit
 RT of replicative DNA polymerase III, from Vibrio cholerae strain
 RT C6706.";
 RL Gene 175:281-283(1996).
 DR EMBL: U30472; LAC44579.1; -.
 FT NON_TER 18 18
 SQ SEQUENCE 18 AA; 2153 MW; 18EBCDAD212842EF CRC64;

Query Match 20.7%; Score 30; DB 2; Length 18;
 Best Local Similarity 63.68; Pred. No. 7.8e+02;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 14 FPKPIVRSITK 24
 || || || ||
 Db 8 FPKPIVLETK 18

RESULT 4
 P87468 PRELIMINARY; PRT; 19 AA.
 ID P87468;
 AC P87468;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE ALPHA-GLOBIN (FRAGMENT).
 OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81001900; PubMed=6250724;
 RA Patient R.K., Elkington J.A., Kay R.M., Williams J.G.:
 RT "Internal organization of the major adult alpha- and beta-globin genes
 RT of X. laevis.";
 RL Cell 21:565-573(1980).
 CC -|- SIMILARITY: TO GLOBIN FAMILY.
 DR EMBL: V01431; CAA24695.1; -.
 DR HSP: P01990; IAAF.
 DR InterPro: IPR000971; Globin.
 DR PROSITE: PS01033; GLOBIN; 1.
 KW Heme; Oxygen transport; Respiratory protein.
 FT NON_TER 1 1
 SQ SEQUENCE 19 AA; 2144 MW; AAC19D73A95C9AC6 CRC64;

Query Match 20.7%; Score 30; DB 13; Length 19;
 Best Local Similarity 54.5%; Pred. No. 8.3e+02;
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 5 YQSRVDHPHP 15
 | ||| : ||
 Db 9 YDLRVDPGNFP 19

RESULT 5
 Q9QV92 PRELIMINARY; PRT; 22 AA.
 AC Q9QV92;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE P93 CALCIUM-BINDING PROTEIN (FRAGMENT).
 OS Rattus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10118;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=33179436; PubMed=8440713;
 RA Gilchrist J.S., Pierce G.N.:
 RT "Identification and purification of a calcium-binding protein in
 RT hepatic nuclear membranes.";
 RL J. Biol. Chem. 268:4291-4299(1993).
 DR InterPro: IPR001580; Calreticulin.
 DR Pfam: PF00262; calreticulin; 1.
 DR ProDom: PD001865; Calreticulin; 1.
 SQ SEQUENCE 22 AA; 2407 MW; B82830E4477E25F9 CRC64;

Query Match 20.7%; Score 30; DB 11; Length 22;
 Best Local Similarity 28.68; Pred. No. 9.5e+02;
 Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 1 CGEGYQSRVDHPHF 14
 || : || : ||
 Db 7 CGVWRPMDPNPY 20

RESULT 6
 Q52918 PRELIMINARY; PRT; 20 AA.
 ID Q52918;
 AC Q52918;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE DELTA-ALA SYNTHETASE GENE 5' REGION (DELTA AMINOLEVULINIC ACID
 DE SYNTHETASE).

OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=85297790; PubMed=2994020;
RA Leong S.A., Williams P.H., Ditta G.S.;
RT "Analysis of the 5' regulatory region of the gene for delta-
aminovaleric acid synthetase of Rhizobium meliloti.";
RL Nucleic Acids Res. 13:5965-5976(1985).
DR EMBL: X02853; CAA26608.1; -
SQ SEQUENCE 20 AA; 2441 MW; BE09CF11ABCC3BB2 CRC64;

Query Match 20.0%; Score 29; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 1.2e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 13 HFPKPIV 19
||| |:
Db 4 HFPSPFL 10

RESULT 7
Q9QSP0 PRELIMINARY; PRT; 23 AA.
ID Q9QSP0
AC Q9QSP0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE L. PROTEIN (FRAGMENT).
GN L.
OS Australian bat lyssavirus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Lyssavirus.
OX NCBI_TaxID=90961;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN-INSECTIVOROUS ISOLATE;
RA Gould A.R., Kattenbelt J.A., Hyatt A.D., Gumley S.G., Lunt R.A.;
RT "Characterisation of a variant of Australian Bat Lyssavirus isolated
from an insectivorous bat and comparison to virus isolates from
Pteropid bats.";
RT Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF081020; AAD47900.1; -
FT NON_TER 23
SQ SEQUENCE 23 AA; 2599 MW; 7B584A2225E76530 CRC64;

Query Match 19.7%; Score 28.5; DB 12; Length 23;
Best Local Similarity 35.0%; Pred. No. 1.7e+03;
Matches 7; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

QY 2 GEGYQSRVDHPHFKPIVRS 21
||| |:
Db 5 GEVYDDPID-PVEPELKT 23

RESULT 8
Q9P2A2 PRELIMINARY; PRT; 14 AA.
ID Q9P2A2
AC Q9P2A2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE TRUNCATED ALDO-KETO REDUCTASE (FRAGMENT).
GN TRUNCATED AKR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RT TISSUE=LIVER;
RX MEDLINE=20138537; PubMed=10672042;
RA Nishizawa M., Nakajima T., Yasuda K., Kanzaki H., Sasaguri Y.,
RA Watanabe K., Ito S.;
RT "Close kinship of human 20alpha-hydroxysteroid dehydrogenase gene with
three aldo-keto reductase genes.";
RL Genes Cells 5:111-125(2000).
DR EMBL: AB037903; BAA92888.1; -
FT NON_TER 1
SQ SEQUENCE 14 AA; 1632 MW; 47EB1EE28D59A8D7 CRC64;

Query Match 19.3%; Score 28; DB 4; Length 14;
Best Local Similarity 56.7%; Pred. No. 1.2e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 13 HFPKPI 18
||| |:
Db 2 HFPVPL 7

RESULT 9
Q44041 PRELIMINARY; PRT; 21 AA.
ID Q44041
AC Q44041; Q44040;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUL-1997 (TrEMBLrel. 04, Last annotation update)
DE HYPOTHETICAL 2.6 KDA PROTEIN.
OS Anabaena flos-aquae.
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
OX NCBI_TaxID=1166;
RN [1]
RN SEQUENCE FROM N.A.
RA Hayes P.K.;
RL Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.
DR EMBL: M32060; AAA82498.1; -
DR EMBL: M32060; AAA82496.1; -
KW Hypothetical protein.
SQ SEQUENCE 21 AA; 2611 MW; 0E4C03BE93E16532 CRC64;

Query Match 19.3%; Score 28; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PHFP 15
||| |:
Db 9 PHFP 12

RESULT 10
Q85496 PRELIMINARY; PRT; 23 AA.
ID Q85496
AC Q85496;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE ENVELOPE PROTEIN (FRAGMENT).
OS Bovine leukemia virus (BLV).
OC Viruses; Retroviridae; Retroviridae; Deltaretrovirus.
OX NCBI_TaxID=11901;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=87198886; PubMed=3032284;
RX Rice N.R., Simex S.L., Dubois G.C., Showalter S.D., Gilden R.V.,
RA Stephens R.M.;
RT "Expression of the bovine leukemia virus X region in virus-infected
cells.";
RL J. Virol. 61:1577-1585(1987).
DR EMBL: M16017; AAA87336.1; -
KW Envelope protein.

```
FT NON_TER 23
SQ SEQUENCE 23 AA; 2877 MW; F85E7CBED5440B08 CRC64;

Query Match 19.3%; Score 28; DB 12; Length 23;
Best Local Similarity 44.4%; Pred. No. 2e+03;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 15 PKPIVRSIT 23
   I::I::
Db 11 PQPIRWVS 19

RESULT 11
Q9UKK1 PRELIMINARY; PRT; 25 AA.
AC Q9UKK1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE METALLOPROTEINASE-DISINTEGRIN (FRAGMENT).
GN ADAM23.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99453762; PubMed=10524237;
RA Polidexter K., Nelson N., DuBoise R.F., Black R.A., Cerretti D.P.;
RT "Identification of seven metalloproteinase-disintegrin (ADAM) genes
   from genomic libraries.";
RL Gene 237:61-70(1999).
DR EMBL; AF158641; AAD5252.1; -.
KW Integrin.
FT NON_TER 25
FT NON_TER 25
SQ SEQUENCE 25 AA; 2705 MW; 818B03B30EF0A19C CRC64;

Query Match 19.3%; Score 28; DB 4; Length 25;
Best Local Similarity 80.0%; Pred. No. 2.2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEGY 5
   I::I
Db 7 CGNGY 11

RESULT 12
Q9BM01 PRELIMINARY; PRT; 25 AA.
AC Q9BM01;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE MARINER-LIKE TRANSPOSASE (FRAGMENT).
OS Hydra littoralis.
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=42240;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON=MARI;
RX MEDLINE=20570504; PubMed=11121049;
RA Arkhipova I., Meselson M.;
RT "Transposable elements in sexual and asexual taxa.";
RL proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
DR EMBL; AY014005; AAG59977.1; -.
FT NON_TER 1
FT NON_TER 25
SQ SEQUENCE 25 AA; 2777 MW; CBF9DD69D7AD074C CRC64;

Query Match 18.6%; Score 27; DB 5; Length 11;
Best Local Similarity 80.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEGY 5
   I::I
Db 11

Query Match 19.3%; Score 28; DB 5; Length 25;
Best Local Similarity 36.4%; Pred. No. 2.2e+03;
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 12 PHFKPIVRSI 22
   I::I::
Db 3 PHVAQPTLRKL 13

RESULT 13
Q9TRY2 PRELIMINARY; PRT; 25 AA.
AC Q9TRY2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE INSULIN-LIKE GROWTH FACTOR-BINDING PROTEIN-6, IGFBP-6.
OS Sus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9826;
RN [1]
RP SEQUENCE.
RX MEDLINE=92049376; PubMed=1719383;
RA Shimasaki S., Gao L., Shimonaka M., Ling N.;
RT "Isolation and molecular cloning of insulin-like growth factor-binding
   protein-6.";
RL Mol. Endocrinol. 5:938-948(1991).
DR InterPro; IPR000867; IGFBP.
DR Pfam; PF00219; IGFBP; 1.
SQ SEQUENCE 25 AA; 2310 MW; 4077663151E71212 CRC64;

Query Match 19.3%; Score 28; DB 6; Length 25;
Best Local Similarity 57.1%; Pred. No. 2.2e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEGYQS 7
   I::I::
Db 6 CGGQVQT 12

RESULT 14
Q9NL65 PRELIMINARY; PRT; 11 AA.
AC Q9NL65;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ASABF-DELTA (FRAGMENT).
GN ASABF-DELTA.
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE FROM N.A.
RA Kato Y.;
RT "Ascaris suum asabf-delta gene, exon 2.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029815; BAA89496.1; -.
FT NON_TER 1
FT NON_TER 11
SQ SEQUENCE 11 AA; 1187 MW; 8BADDD0CD1EAB5861 CRC64;

Query Match 18.6%; Score 27; DB 5; Length 11;
Best Local Similarity 80.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEGY 5
   I::I
Db 11
```

Db 2 CGTGY 6

RESULT 15

Q9UCC0

ID Q9UCC0 PRELIMINARY; PRT; 15 AA.

AC Q9UCC0;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE INSULIN-LIKE GROWTH FACTOR RECEPTOR ALPHA SUBUNIT (FRAGMENT).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE.

RX MEDLINE=94079885; PubMed=8257688;

RA Kasuya J., Paz I.B., Maddux B.A., Goldfine I.D., Hefta S.A.,

RA Fujita-Yamaguchi Y.;

RT "Characterization of human placental insulin-like growth factor-

RT I/insulin hybrid receptors by protein microsequencing and

RT purification.";

RL Biochemistry 32:13531-13536(1993).

SQ SEQUENCE 15 AA; 1721 MW; 98BC151D6D81784B CRC64;

Query Match 18.6%; Score 27; DB 4; Length 15;

Best Local Similarity 45.5%; Pred. No. 1.8e+03;

Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDH 11

|| | | |

Db 3 CGPGIDIRNDY 13

Search completed: March 4, 2002, 13:21:52

Job time: 653 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:10:52 ; Search time 37.41 Seconds
(without alignments)
15.038 Million cell updates/sec

Title: US-09-701-623C-7

Perfect score: 145

Sequence: 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 123821

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1 | 103 | 71.0 | 25 | 3 US-09-100-414B-95 | Sequence 95, Appl |
| 2 | 103 | 71.0 | 25 | 4 US-09-303-323-95 | Sequence 95, Appl |
| 3 | 62 | 42.8 | 22 | 2 US-08-232-539D-19 | Sequence 19, Appl |
| 4 | 62 | 42.8 | 24 | 2 US-08-232-539D-20 | Sequence 20, Appl |
| 5 | 37 | 25.5 | 15 | 2 US-08-630-645-7 | Sequence 7, Appl |
| 6 | 37 | 25.5 | 15 | 5 PCT-US96-10220-7 | Sequence 7, Appl |
| 7 | 35 | 24.1 | 13 | 6 517197-5 | Patent No. 517197 |
| 8 | 35 | 24.1 | 20 | 2 US-08-564-972-54 | Sequence 54, Appl |
| 9 | 35 | 24.1 | 20 | 2 US-08-564-972-55 | Sequence 55, Appl |
| 10 | 34 | 23.4 | 13 | 4 US-09-258-754-218 | Sequence 218, App |
| 11 | 34 | 23.4 | 13 | 4 US-09-042-107-218 | Sequence 218, App |
| 12 | 34 | 23.4 | 18 | 4 US-08-602-999A-371 | Sequence 371, App |
| 13 | 34 | 23.4 | 20 | 1 US-07-678-974D-5 | Sequence 5, Appl |
| 14 | 34 | 23.4 | 20 | 2 US-08-945-168-10 | Sequence 10, Appl |
| 15 | 33 | 22.8 | 17 | 4 US-08-602-999A-398 | Sequence 398, App |
| 16 | 33 | 22.8 | 17 | 4 US-08-602-999A-399 | Sequence 399, App |
| 17 | 33 | 22.8 | 18 | 4 US-09-177-249-298 | Sequence 298, App |
| 18 | 32 | 22.1 | 18 | 4 US-09-461-697-30 | Sequence 30, Appl |
| 19 | 32 | 22.1 | 18 | 5 PCT-US93-03748-8 | Sequence 8, Appl |
| 20 | 32 | 22.1 | 20 | 1 US-07-678-974D-7 | Sequence 7, Appl |
| 21 | 32 | 22.1 | 20 | 2 US-08-945-168-12 | Sequence 12, Appl |
| 22 | 32 | 22.1 | 20 | 4 US-08-602-999A-120 | Sequence 120, App |
| 23 | 31 | 21.4 | 11 | 2 US-08-630-645-8 | Sequence 8, Appl |
| 24 | 31 | 21.4 | 11 | 5 PCT-US96-10220-8 | Sequence 8, Appl |
| 25 | 31 | 21.4 | 16 | 1 US-08-447-010-14 | Sequence 14, Appl |
| 26 | 31 | 21.4 | 18 | 1 US-08-395-602A-1 | Sequence 1, Appl |
| 27 | 31 | 21.4 | 18 | 2 US-08-021-625D-1 | Sequence 1, Appl |

| | | | | | |
|----|----|------|----|---------------------|-------------------|
| 28 | 31 | 21.4 | 19 | 1 US-08-116-733-12 | Sequence 12, Appl |
| 29 | 31 | 21.4 | 19 | 1 US-08-469-615-9 | Sequence 9, Appl |
| 30 | 31 | 21.4 | 19 | 1 US-08-466-763-9 | Sequence 9, Appl |
| 31 | 31 | 21.4 | 19 | 2 US-08-411-142A-9 | Sequence 9, Appl |
| 32 | 31 | 21.4 | 20 | 4 US-08-986-659B-8 | Sequence 8, Appl |
| 33 | 31 | 21.4 | 23 | 1 US-08-395-602A-2 | Sequence 2, Appl |
| 34 | 31 | 21.4 | 23 | 2 US-08-021-625D-2 | Sequence 2, Appl |
| 35 | 31 | 21.4 | 23 | 4 US-08-986-659B-9 | Sequence 9, Appl |
| 36 | 31 | 21.4 | 23 | 4 US-08-986-659B-33 | Sequence 33, Appl |
| 37 | 30 | 20.7 | 10 | 1 US-07-801-812A-8 | Sequence 8, Appl |
| 38 | 30 | 20.7 | 10 | 1 US-08-487-568-8 | Sequence 8, Appl |
| 39 | 30 | 20.7 | 12 | 1 US-08-433-318A-86 | Sequence 86, Appl |
| 40 | 30 | 20.7 | 12 | 2 US-08-922-048-86 | Sequence 86, Appl |
| 41 | 30 | 20.7 | 12 | 2 US-08-323-686-15 | Sequence 15, Appl |
| 42 | 30 | 20.7 | 12 | 5 PCT-US95-11127-8 | Sequence 8, Appl |
| 43 | 30 | 20.7 | 12 | 5 PCT-US96-06270-86 | Sequence 86, Appl |
| 44 | 30 | 20.7 | 16 | 2 US-08-484-905-43 | Sequence 43, Appl |
| 45 | 30 | 20.7 | 16 | 3 US-08-481-985B-43 | Sequence 43, Appl |

ALIGNMENTS

RESULT 1
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-95

Query Match 71.0%; Score 103; DB 3; Length 25;
Best Local Similarity 68.0%; Pred. No. 1.5e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Db 1 CGETYSRVTHPHLPALMRSTTKC 25

RESULT 2

US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10151-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 71.0%; Score 103; DB 4; Length 25;
Best Local Similarity 68.0%; Pred. No. 1.5e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | :|| |||
Db 1 CGETYQSRVTHPLPRALMRSTKC 25

RESULT 3

US-08-232-539D-19
; Sequence 19, Application US/08232539D
; Patent No. 5965709
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; TITLE OF INVENTION: Ige Antagonists
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,539D
FILING DATE: 21-Apr-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/178583
FILING DATE: 07-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/744768
FILING DATE: 14-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Svoboda, Craig G.
REGISTRATION NUMBER: 39,044
REFERENCE/DOCKET NUMBER: P0718P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-1489
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-232-539D-19

Query Match 42.8%; Score 62; DB 2; Length 22;
Best Local Similarity 55.8%; Pred. No. 0.0028;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIV 19
||| ||||| ||| | :|| |||
Db 5 GETYQCRVTHPLPRALM 22

RESULT 4

US-08-232-539D-20
; Sequence 20, Application US/08232539D
; Patent No. 5965709
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; TITLE OF INVENTION: Ige Antagonists
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881

Db 1 SRGDLPPFPVPI 12

```

; LENGTH: 13
; 1 2 3 4 5 6 7 8 9 10 11 12 13

```

LENG
; 11107
F

Query Match 24.1%; Score 35; DB 6; Length 13;
Best Local Similarity 83.3%; Pred. No. 23;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGEGYQ 6
DB 3 CGQGYQ 8

RESULT 8
US-08-564-972-54
; Sequence 54, Application US/08564972
; Patent No. 5843462
; GENERAL INFORMATION:
; APPLICANT: Conti-Fine, B. M.
; TITLE OF INVENTION: DIPHTHERIA TOXIM EPITOPES
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/564,972
; FILING DATE: 30-NOV-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 600.344US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-564-972-54

Query Match 24.1%; Score 35; DB 2; Length 20;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GYQSRVDH 11
DB 13 GYQKTVDH 20

RESULT 9
US-08-564-972-55
; Sequence 55, Application US/08564972
; Patent No. 5843462
; GENERAL INFORMATION:

; APPLICANT: Conti-Fine, B. M.
; TITLE OF INVENTION: DIPHTHERIA TOXIM EPITOPES
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/564,972
; FILING DATE: 30-NOV-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 600.344US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-339-0331
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-564-972-55

Query Match 24.1%; Score 35; DB 2; Length 20;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 GYQSRVDH 11
DB 3 GYQKTVDH 10

RESULT 10
US-09-258-754-218
; Sequence 218, Application US/09258754
; Patent No. 6174687
; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; APPLICANT: Rajotte, Daniel
; TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using
; TITLE OF INVENTION: Membrane Dipeptidase
; FILE REFERENCE: P-LJ 3443
; CURRENT APPLICATION NUMBER: US/09/258,754
; CURRENT FILING DATE: 1999-02-26
; EARLIER APPLICATION NUMBER: 09/042,107
; EARLIER FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 452
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 218
; LENGTH: 13
; TYPE: PRT

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-258-754-218

Query Match 23.4%; Score 34; DB 4; Length 13;
Best Local Similarity 62.5%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEGYQR 8
|||
DB 1 CGPGYQAQ 8

RESULT 11
US-09-107-218
; Sequence 218, Application US/09042107
; Patent No. 6232287

; GENERAL INFORMATION:
; APPLICANT: Ruoslahti, Erkki
; APPLICANT: Pasqualini, Renata
; TITLE OF INVENTION: Molecules that Home to Various Selected Organs or
; TITLE OF INVENTION: Tissues
; FILE REFERENCE: P-LJ 2892
; CURRENT APPLICATION NUMBER: US/09/042,107
; CURRENT FILING DATE: 1998-03-13
; NUMBER OF SEQ ID NOS: 436
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 218
; LENGTH: 13
; TYPE: PRT

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-042-107-218

Query Match 23.4%; Score 34; DB 4; Length 13;
Best Local Similarity 62.5%; Pred. No. 33;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGEGYQR 8
|||
DB 1 CGPGYQAQ 8

RESULT 12
US-08-602-999A-371
; Sequence 371, Application US/08602999A
; Patent No. 6184205

; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999A
; FILING DATE: 16-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 371:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-602-999A-371

Query Match 23.4%; Score 34; DB 4; Length 18;
Best Local Similarity 71.4%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 11 HPHFKRP 17
|||||
DB 1 HPHAPPP 7

RESULT 13

US-07-678-974D-5
; Sequence 5, Application US/07678974D
; Patent No. 5629146

; GENERAL INFORMATION:
; APPLICANT: DILLNER, JOAKIM
; APPLICANT: DILLNER, LENA
; TITLE OF INVENTION:
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 67
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERMAN & AISENBERG
; STREET: 1730 RHODE ISLAND AVENUE, N.W.,
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-3186

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/678,974D
; FILING DATE: 25-JUN-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: AISENBERG, Irwin M.
; REGISTRATION NUMBER: 19,007
; REFERENCE/DOCKET NUMBER: SG19171
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-293-1404
; TELEFAX: 202-872-0493
; TELEX: 440 069 AIS UI

; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-07-678-974D-5

Query Match: 23.4%; Score 34; DB 1; Length 20;
Best Local Similarity 63.6%; Pred. No. 52;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 9 VDHPPFP--KP 17
| ||||| |
Db 10 VGHPPFPFKP 20

RESULT 14
US-08-945-168-10
; Sequence 10, Application US/08945168
; Patent No. 5989548
; GENERAL INFORMATION:
; APPLICANT: DILLNER, JOAKIM
; TITLE OF INVENTION: PEPTIDE-BASED VACCINE AGAINST PAPILLOMA
; TITLE OF INVENTION: VIRUS
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: F.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/945,168
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/SE96/00533
; FILING DATE: 23-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: SE 9501512-9
; FILING DATE: 24-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 7752-0002-0 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-945-168-10

Query Match: 23.4%; Score 34; DB 2; Length 20;
Best Local Similarity 63.6%; Pred. No. 52;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 9 VDHPPFP--KP 17
| ||||| |
Db 10 VGHPPFPFKP 20

RESULT 15
US-08-602-999A-398
; Sequence 398, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.

; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: OULLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999A
; FILING DATE: 16-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 398:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-602-999A-398

Query Match: 22.8%; Score 33; DB 4; Length 17;
Best Local Similarity 71.4%; Pred. No. 63;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 HPHPFKP 17
| ||||| |
Db 4 HPNFKP 10

Search completed: March 4, 2002, 13:10:52
Job time: 299 sec

THIS PAGE BLANK (USPTO,

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:11:39 ; Search time 40.99 Seconds
(without alignments)
46.459 Million cell updates/sec

Title: US-09-701-623C-8
Perfect score: 140
Sequence: 1 CGYGQSVIVDRPDPFPKIVRSITLC 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 4959

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|--------|---------------------|
| 1 | 37 | 26.4 | 23 | S65374 | cytochrome-c oxida |
| 2 | 33 | 23.6 | 18 | B49254 | Tcr C gamma 1 chai |
| 3 | 30 | 21.4 | 15 | PS0450 | 23K protein 4307 |
| 4 | 30 | 21.4 | 20 | S44465 | pyrrhocoriclin - Py |
| 5 | 30 | 21.4 | 22 | S73389 | hypothetical prote |
| 6 | 29 | 20.7 | 17 | D48138 | d(TTAGGG)n-binding |
| 7 | 28.5 | 20.4 | 23 | E39855 | paralytic peptide |
| 8 | 28.5 | 20.4 | 23 | D39855 | paralytic peptide |
| 9 | 28 | 20.0 | 13 | S21152 | tryptophyllin-rela |
| 10 | 28 | 20.0 | 17 | S33609 | extensin - maize (|
| 11 | 28 | 20.0 | 22 | S52357 | hypothetical prote |
| 12 | 28 | 20.0 | 23 | A32473 | histidine-rich pro |
| 13 | 28 | 20.0 | 25 | JP0045 | ribosomal protein |
| 14 | 27 | 19.3 | 21 | S47207 | T-cell receptor J- |
| 15 | 27 | 19.3 | 25 | B57001 | endo-1,4-beta-xyla |
| 16 | 26.5 | 18.9 | 24 | T42257 | phosphoprotein pho |
| 17 | 26 | 18.6 | 11 | PN0042 | statmin - mouse (|
| 18 | 26 | 18.6 | 13 | A05174 | tryptophyllin-l3 |
| 19 | 26 | 18.6 | 15 | PA0088 | protein QF200051 - |
| 20 | 26 | 18.6 | 22 | JN0910 | M4-(beta-N-acetyl |
| 21 | 26 | 18.6 | 23 | C39855 | paralytic peptide |
| 22 | 26 | 18.6 | 23 | B61079 | listeriolysin O - |
| 23 | 26 | 18.6 | 25 | I40692 | cenA protein (Iga1 |
| 24 | 25 | 17.9 | 14 | S50900 | chlorophyll a/b-bi |
| 25 | 25 | 17.9 | 19 | S59485 | hydroxyproline-rich |
| 26 | 25 | 17.9 | 22 | S55236 | T669 kinase - huma |
| 27 | 25 | 17.9 | 23 | S37452 | gene E6 protein (c |
| 28 | 25 | 17.9 | 23 | A56357 | tyrosine kinase su |
| 29 | 25 | 17.9 | 24 | S53793 | actin - mouse (fra |

RESULT 1
S65374
cytochrome-c oxidase (EC 1.9.3.1) chain IV, cardiac - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 16-Jul-1999
C:Accession: S65374
R:Schaeffer, H.; Noack, H.; Halandk, W.; Brandt, U.; von Jagow, G.
Eur. J. Biochem. 230, 235-241, 1995
A:Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-t
A:Reference number: S65372; MUID:95324529
A:Accession: S65374
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-23 <SCH>
C:Superfamily: cytochrome-c oxidase chain IV
C:Keywords: cardiac muscle; heart; oxidoreductase

ALIGNMENTS

| | | | | | | |
|----|------|------|----|---|--------|--------------------|
| 30 | 24.5 | 17.5 | 23 | 2 | A39855 | paralytic peptide |
| 31 | 24 | 17.1 | 14 | 2 | P0026 | calotropin DI - mu |
| 32 | 24 | 17.1 | 16 | 2 | S30384 | hypothetical prote |
| 33 | 24 | 17.1 | 16 | 2 | A20190 | hypodermin B - ear |
| 34 | 24 | 17.1 | 17 | 2 | JP0046 | ribosomal protein |
| 35 | 24 | 17.1 | 18 | 2 | S59490 | translation elonga |
| 36 | 24 | 17.1 | 20 | 2 | JP0050 | ribosomal protein |
| 37 | 24 | 17.1 | 20 | 2 | JP0051 | ribosomal protein |
| 38 | 24 | 17.1 | 20 | 2 | PH1341 | Ig heavy chain DJ |
| 39 | 24 | 17.1 | 20 | 2 | PQ0071 | T-cell receptor be |
| 40 | 24 | 17.1 | 25 | 2 | JP0047 | ribosomal protein |
| 41 | 24 | 17.1 | 25 | 2 | JP0048 | ribosomal protein |
| 42 | 23.5 | 16.8 | 14 | 2 | E33098 | 214K exoantigen (v |
| 43 | 23.5 | 16.8 | 19 | 2 | E49048 | T-cell receptor be |
| 44 | 23.5 | 16.8 | 24 | 2 | T42441 | protein phosphatas |
| 45 | 23 | 16.4 | 10 | 2 | C45474 | thrombospondin 2 - |

Query Match 26.4%; Score 37; DB 2; Length 23;
Best Local Similarity 53.8%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 YGYQSVIVDRPDPFP 15
| | | | |
DB 11 YALPSYVDRRDYP 23

RESULT 2
B49254
Tcr C gamma 1 chain V-J region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 23-Jul-1999
C:Accession: B49254
R:Ezquerria, A.; Wilde, D.B.; McConnell, T.J.; Sturmhofel, K.; Valas, R.B.; Shevach, E
Eur. J. Immunol. 22, 491-498, 1992
A:Title: Mouse autoreactive gamma/delta T cells. II. Molecular characterization of th
A:Reference number: A49037; MUID:92164730
A:Accession: B49254
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-18 <EZQ>
A:Cross-references: GB:S90688; NID:g246308; PIDN:AAB21557.1; PID:g246309
A:Experimental source: Y93A cells
A>Note: sequence extracted from NCBI backbone (NCBIN:90688, NCBTP:90690)
C:Superfamily: immunoglobulin V region; immunoglobulin homology

Query Match 23.6%; Score 33; DB 2; Length 18;
Best Local Similarity 71.4%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYGQSV 7

Db 5 CSYGYS 11
| | | | |

RESULT 3

PS0450

23K protein 4307 - rice (strain Nihonbare) (fragment)
C:Species: Oryza sativa (rice)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 11-Apr-1995

C:Accession: PS0450

R:Tsugeta, A.; Miyatake, N.

submitted to JIPID, April 1993

A:Reference number: PS0208

A:Accession: PS0450

A:Molecule type: protein

A:Residues: 1-15 <TSU>

A:Experimental source: germ, strain Nihonbare

C:Comment: molecular weight 23K, pi 5.3.

Query Match 21.4%; Score 30; DB 2; Length 15;

Best Local Similarity 66.7%; Pred. No. 2.2e+02;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 5 YQSVDRPD 13

| | | | |

Db 2 YQDIPDRKD 10

RESULT 4

S44465

pyrrolicorin - Pyrrhocoris apterus

C:Species: Pyrrhocoris apterus

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999

C:Accession: S44465

R:Cociancich, S.; Dupont, A.; Hegy, G.; Lanot, R.; Holder, F.; Hetru, C.; Hoffmann, J.A.

Biochem. J. 300, 567-575, 1994

A:Title: Novel inducible antibacterial peptides from a hemipteran insect, the sap-sucking

A:Reference number: S44463; MUID:94271176

A:Accession: S44465

A:Molecule type: protein

A:Residues: 1-20 <CCC>

C:Function:

A:Description: antibacterial protein

A:Note: active against Gram-negative bacteria

C:Keywords: antibacterial; hemolymph; immune response

Query Match 21.4%; Score 30; DB 2; Length 20;

Best Local Similarity 50.0%; Pred. No. 3e+02;

Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 7 SIVDRPDPPKPI 18

| | | | |

Db 5 SYLPRPTPPRPI 16

RESULT 5

S73389

hypothetical protein A05_orf139 - Mycoplasma pneumoniae (strain ATCC 29342)

C:Species: Mycoplasma pneumoniae

A:Variety: ATCC 29342

C:Date: 26-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999

C:Accession: S73389

R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkel, E.; Li, B.C.; Herrmann, R.

Nucleic Acids Res. 24, 4420-4449, 1996

A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae

A:Reference number: S73327; MUID:97105885

A:Accession: S73389

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-22 <HIM>

A:Cross-references: EMBL:AE000008; GB:U000089; NID:g1673711; PIDN:AAB95715.1; PID:g1673711

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetics:
A:Genetic code: SGC3
A:Start codon: GTG

Query Match 21.4%; Score 30; DB 2; Length 22;
Best Local Similarity 50.0%; Pred. No. 3.4e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

Qy 2 GYGYSIVDR--PDPP 15

| | | | |

Db 6 GSPYHWNLDLRNPDP 21

RESULT 6

D48138

d(TTAGG)n-binding protein B37 - human (fragment)

N:Alternate names: type A-B heterogeneous nuclear ribonucleoprotein homolog

C:Species: Homo sapiens (man)

C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 28-Feb-1997

C:Accession: D48138

R:Ishikawa, F.; Matunis, M.J.; Dreyfuss, G.; Cech, T.R.

Mol. Cell. Biol. 13, 4301-4310, 1993

A:Title: Nuclear proteins that bind the pre-mRNA 3' splice site sequence r(UUAG/G) an

A:Reference number: A48138; MUID:93309464

A:Accession: D48138

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-17 <ISH>

A:Experimental source: HeLa cell nuclei

A:Note: sequence extracted from NCBI backbone (NCBIP:134642)

Query Match 20.7%; Score 29; DB 2; Length 17;

Best Local Similarity 38.5%; Pred. No. 3.6e+02;

Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 5 YQSVDRPDPPKPI 17

| | | | |

Db 1 FGEVVDXTDTPDP 13

RESULT 7

E39855

paralytic peptide III - beet armyworm

C:Species: Spodoptera exigua (beet armyworm)

C:Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 30-Sep-1993

C:Accession: E39855

R:Skinner, W.S.; Dennis, P.A.; Li, J.P.; Summerfelt, R.M.; Carney, R.L.; Quistad, G.B.

J. Biol. Chem. 266, 12873-12877, 1991

A:Title: Isolation and identification of paralytic peptides from hemolymph of the lep

A:Reference number: A39855; MUID:91302298

A:Accession: E39855

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-23 <SKI>

C:Superfamily: paralytic peptide I

Query Match 20.4%; Score 28.5; DB 2; Length 23;

Best Local Similarity 41.2%; Pred. No. 6e+02;

Matches 7; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

Qy 1 CGYGYQSIVD---RPDF 14

| | | | |

Db 7 CTPGYQRTADGRCKPTF 23

RESULT 8

D39855

paralytic peptide II - beet armyworm

C:Species: Spodoptera exigua (beet armyworm)

QY 2 GYGYSIVDRPDF--PKP 17
 |||| | : |||
 Db 3 GGY-----GPPYTPTPKP 16

```
Query Match      20.0%; Score 28; DB 2; Length 25;
Best Local Similarity 23.5%; Pred. No. 7.9e+02;
Matches 4: Conservative 7; Mismatches 6; Indels 0; Gaps 0;
```

Qy 6 QSVDRDPFPRKPIVRSI 22
:|::||: :|::
Db 9 RSVIGRPETQRKTVEAL 25

RESULT 14

S47207
T-cell receptor J-alpha wNVIII.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 05-Nov-1999
C:Accession: S47207
R:Plaza, A.; Kono, E.H.; Theofilopoulos, A.N.
submitted to the EMBL Data Library, February 1993
A:Reference number: S40133
A:Accession: S47207
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-21 <PLA>
A:Cross-references: EMBL:X71032; NID:g506610; PIDN:CAA50349.1; PID:g510318
C:Keywords: T-cell receptor

Query Match 19.3%; Score 27; DB 2; Length 21;
Best Local Similarity 55.6%; Pred. No. 9.2e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 7 SIVDRPDPF 15
|:| ||| |
Db 10 SLVPEPDCP 18

RESULT 15

B57001
endo-1,4-beta-xylanase (EC 3.2.1.8) 2 - Streptomyces roseiscleroticus (fragment)
C:Species: Streptomyces roseiscleroticus
C:Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 06-Dec-1996
C:Accession: B57001
R:Grabski, A.C.; Forrester, I.T.; Patel, R.; Jeffries, T.W.
Protein Expr. Purif. 4, 120-129, 1993
A:Title: Characterization and N-terminal amino acid sequences of beta-(1-4)endoxylanases
A:Reference number: A57001; MUID:93229899
A:Accession: B57001
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-25 <GRA>
A:Experimental source: strain NRRL B-11019
A>Note: sequence extracted from NCBI backbone (NCBIP:130008)

C:Function:
A:Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
A:Pathway: xylan degradation
C:Superfamily: endo-1,4-beta-xylanase; endo-1,4-beta-xylanase homology
C:Keywords: extracellular protein; glycosidase; hydrolase; polysaccharide degradation

Query Match 19.3%; Score 27; DB 2; Length 25;
Best Local Similarity 41.7%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GYGQSIIVDRPD 13
||| | |
Db 14 GYYYSFXTDAFN 25

Search completed: March 4, 2002, 13:11:40
Job time: 292 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:20:28 ; Search time 24.21 Seconds
(without alignments)
37.861 Million cell updates/sec

Title: US-09-701-623C-8
Perfect score: 140
Sequence: 1 CGYGYQSIVDRPDKPIVRSITLC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues
Total number of hits satisfying chosen parameters: 1446

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 35.5 | 25.4 | 16 | 1 MK2B_PALPR | P80410 palomena pr |
| 2 | 33.5 | 23.9 | 15 | 1 MK1_PALPR | P80408 palomena pr |
| 3 | 33.5 | 23.9 | 15 | 1 MK2A_PALPR | P80409 palomena pr |
| 4 | 33.5 | 23.9 | 16 | 1 MK3_PALPR | P80411 palomena pr |
| 5 | 30 | 21.4 | 20 | 1 PYRR_PVRAP | P37362 pyrrhocoris |
| 6 | 28.5 | 20.4 | 23 | 1 CP23_SPOER | P56683 spodoptera |
| 7 | 28.5 | 20.4 | 23 | 1 PAP2_SPOEX | P30256 spodoptera |
| 8 | 28.5 | 20.4 | 23 | 1 PAP3_SPOEX | P30257 spodoptera |
| 9 | 26 | 18.6 | 23 | 1 TP13_PHYRO | P04096 phyllomedusa |
| 10 | 26 | 18.6 | 23 | 1 PAP1_SPOEX | P30255 spodoptera |
| 11 | 26 | 18.6 | 24 | 1 CAMT_PINPS | P81081 pinus pinas |
| 12 | 24.5 | 17.5 | 23 | 1 PAP1_MANSE | P30253 manduca sex |
| 13 | 24 | 17.1 | 14 | 1 CALL_CALGI | P20728 calotropis |
| 14 | 24 | 17.1 | 16 | 1 YMOR_PSEPU | Q02210 pseudomonas |
| 15 | 24 | 17.1 | 19 | 1 UP21_UPEIN | P82027 uperoleia i |
| 16 | 24 | 17.1 | 19 | 1 UP24_UPEIN | P82030 uperoleia i |
| 17 | 24 | 17.1 | 24 | 1 FRE4_LITIN | P82023 litoria inf |
| 18 | 23 | 16.4 | 20 | 1 KORC_METTM | P80906 methanobact |
| 19 | 23 | 16.4 | 20 | 1 UN05_PINPS | P81674 pinus pinas |
| 20 | 22 | 16.4 | 24 | 1 IRBP_SHEEP | P12663 ovis aries |
| 21 | 22 | 15.7 | 11 | 1 BPBP_AKHA | P01021 agkistrodon |
| 22 | 22 | 15.7 | 12 | 1 GRAR_RANRU | P40754 rana rugosa |
| 23 | 22 | 15.7 | 12 | 1 TKNL_KASMA | P08613 kassina mac |
| 24 | 22 | 15.7 | 17 | 1 RUBR_CHLTE | P58025 chlorobium |
| 25 | 22 | 15.7 | 18 | 1 RL23_HALCU | P05975 halobacteri |
| 26 | 22 | 15.7 | 19 | 1 DUCR_STRGP | P36503 streptomyce |
| 27 | 22 | 15.7 | 21 | 1 LPRM_CORDI | P21232 corynebacte |
| 28 | 22 | 15.7 | 21 | 1 RP71_STRSQ | P37046 streptomyce |
| 29 | 22 | 15.7 | 25 | 1 ANDT_ANDAU | P56684 androctonus |
| 30 | 22 | 15.7 | 25 | 1 CXA4_CONPU | P55963 conus purpu |
| 31 | 21.5 | 15.4 | 24 | 1 ACHA_ELEEL | P09688 electrophor |
| 32 | 21.5 | 15.4 | 25 | 1 GBP_APAKA | P22800 apanteles k |
| 33 | 21 | 15.0 | 10 | 1 HTF_TABAT | P14596 tabanus atr |

ALIGNMENTS

RESULT 1

MK2B_PALPR
ID MK2B_PALPR STANDARD; PRT; 16 AA.
AC P80410;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE METALNIKOWIN IIB.
OS Palomena prasina.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomoidea;
OC Pentatomidae; Palomena.
OX NCBI_TaxID=55431;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RA Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
RT "The inducible antibacterial peptides of the hemipteran insect
Palomena prasina. Identification of a unique family of proline-rich
peptides and of a novel insect defensin.";
RL J. Insect Physiol. 42:81-89(1996).
CC -|- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
CC BACTERIA.
CC -|- INDUCTION: BY BACTERIAL INFECTION.
KW Antibiotic; Insect immunity.
SQ SEQUENCE 16 AA; 2040 MW; AF21407D063B9462 CRC64;

Query Match 25.4%; Score 35.5; DB 1; Length 16;
Best Local Similarity 46.7%; Pred. No. 19;
Matches 7; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Qy 9 VDRPDF-PKPIVRSI 22
||:|:|:|:|:|:|:
Db 1 VDRPDYRPRPWPRNM 15

RESULT 2

MK1_PALPR
ID MK1_PALPR STANDARD; PRT; 15 AA.
AC P80408;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE METALNIKOWIN I.
OS Palomena prasina.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomoidea;
OC Pentatomidae; Palomena.
OX NCBI_TaxID=55431;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RA Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;

```

RT "The inducible antibacterial peptides of the hemipteran insect
RT Palomena prasina. Identification of a unique family of proline-rich
RT peptides and of a novel insect defensin."
RL J. Insect Physiol. 42:81-89(1996).
CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
CC BACTERIA.
CC -1- INDUCTION: BY BACTERIAL INFECTION.
KW Antibiotic; Insect immunity.
SQ SEQUENCE 15 AA; 1838 MW; 21407E663CE46299.CRC64;

Query Match 23.9%; Score 33.5; DB 1; Length 15;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 9 VDRPDF-PKP 17
Db 1 VDKPDYRPRP 10

RESULT 3
MK2A_PALPR STANDARD; PRT; 15 AA.
ID MK2A_PALPR
AC P80409;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE METALNIKOWIN IIA.
OS Palomena prasina.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomoidea;
OC Pentatomidae; Palomena.
OC NCBI_TaxID=55431;
OX [1]
RN [1]
RP TISSUE=Hemolymph;
RA Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
RT "The inducible antibacterial peptides of the hemipteran insect
RT Palomena prasina. Identification of a unique family of proline-rich
RT peptides and of a novel insect defensin."
RL J. Insect Physiol. 42:81-89(1996).
CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
CC BACTERIA.
CC -1- INDUCTION: BY BACTERIAL INFECTION.
KW Antibiotic; Insect immunity.
SQ SEQUENCE 15 AA; 1893 MW; 23835D063B946299.CRC64;

Query Match 23.9%; Score 33.5; DB 1; Length 15;
Best Local Similarity 60.0%; Pred. No. 36;
Matches 6; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 9 VDRPDF-PKP 17
Db 1 VDKPDYRPRP 10

RESULT 4
MK3_PALPR STANDARD; PRT; 16 AA.
ID MK3_PALPR
AC P80411;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE METALNIKOWIN III.
OS Palomena prasina.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomoidea;
OC Beta-Omidae; Palomena.
OC NCBI_TaxID=55431;
OX [1]
RN [1]

```

```

RP SEQUENCE.
RC TISSUE=Hemolymph;
RA Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
RT "The inducible antibacterial peptides of the hemipteran insect
RT Palomena prasina. Identification of a unique family of proline-rich
RT peptides and of a novel insect defensin."
RL J. Insect Physiol. 42:81-89(1996).
CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
CC BACTERIA.
CC -1- INDUCTION: BY BACTERIAL INFECTION.
KW Antibiotic; Insect immunity.
SQ SEQUENCE 16 AA; 2024 MW; A9E3835D063B9462.CRC64;

Query Match 23.9%; Score 33.5; DB 1; Length 16;
Best Local Similarity 60.0%; Pred. No. 39;
Matches 6; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

QY 9 VDRPDF-PKP 17
Db 1 VDKPDYRPRP 10

RESULT 5
PYRR_PYRAP STANDARD; PRT; 20 AA.
ID PYRR_PYRAP
AC P37362; P80307;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE PYRRHOCORICIN.
OS Pyrrhocoris apterus (Sap sucking bug).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC Heteroptera; Panheteroptera; Pentatomomorpha; Lygaeoidea;
OC Pyrrhocoridae; Pyrrhocoris.
OC NCBI_TaxID=37000;
OX [1]
RN [1]
RP TISSUE=Hemolymph;
RA Cociancich S., Dupont A., Hegy G., Lanot R., Holder F., Hetru C.,
RT Hoffmann J. A., Bulet P.;
RT "Novel inducible antibacterial peptides from a hemipteran insect, the
RT sap-sucking bug Pyrrhocoris apterus."
RL Biochem. J. 300:567-575(1994).
RN [2]
RP CARBOHYDRATE-LINKAGE SITE THR-11.
RX MEDLINE=99177428; PubMed=10076062;
RA Hoffmann R., Bulet P., Urge L., Otvoes L. Jr.;
RT "Range of activity and metabolic stability of synthetic antibacterial
RT glycopeptides from insects."
RL Biochim. Biophys. Acta 1426:459-467(1999).
CC -1- FUNCTION: ANTIBACTERIAL PEPTIDE. AFFECTS GRAM-NEGATIVE BACTERIA
CC E. COLI 1106, P. AERUGINOSA, E. COLI D22 AND E. CLOACAE AND
CC GRAM-POSITIVE BACTERIA M. LUTEUS AND B. SUBTILIS.
CC -1- PTM: O-LINKED GLYCAN CONSISTS OF A GAL-GALNAC DISACCHARIDE, O-
CC GLYCOSYLATION IS ESSENTIAL FOR FULL BIOLOGICAL ACTIVITY.
CC -1- SIMILARITY: TO DROSOPHILA DROSOCIN.
DR PIR; S44465; S44465.
KW Antibiotic; Glycoprotein; Insect immunity; Hemolymph.
FT CARBOHYD 11 O-LINKED (GALNAC...).
SQ SEQUENCE 20 AA; 2341 MW; F4320EC2F29462C.CRC64;

Query Match 21.4%; Score 30; DB 1; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 7 SIVDPRDPFKPI 18
Db 5 SYLPRPTPRPI 16

```

```

RESULT 6
CP23_SPOER
ID CP23_SPOER STANDARD; PRT; 23 AA.
AC P56683;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CARDIOACTIVE PEPTIDE CAP23.
OS Spodoptera eridania (Southern armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
OX NCBI_TaxID=37547;
RN [1]
SEQUENCE.
RX MEDLINE=99196260; PubMed=10098624;
RA Furuya K., Hackett M., Cirelli M.A., Schegg K.M., Wang H.,
RA Shabanowitz J., Hunt D.F., Schooley D.A.;
RT "A cardioactive peptide from the southern armyworm, Spodoptera
RT eridania.";
RL Peptides 20:53-61(1999).
CC -1- FUNCTION: HAS EXCITATORY EFFECTS ON A SEMI-ISOLATED HEART FROM
CC LARVAL MANDUCA SEXTA, CAUSING AN INOTROPIC EFFECT AT LOW
CC CONCENTRATIONS OF PEPTIDE AND CHRONOTROPIC AND INOTROPIC EFFECTS
CC AT HIGH DOSES.
CC -1- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
DR InterPro: IPR003463; GBP_PSP.
DR Pfam: PF02425; GBP_PSP; 1.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2519 MW; 0A96D72A70855AE0 CRC64;

Query Match 20.4%; Score 28.5; DB 1; Length 23;
Best Local Similarity 41.2%; Pred. No. 3.3e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 CGYGYQSIVD---RPDF 14
I III I :I I
Db 7 CTFGYQRTADGRCKPTF 23

RESULT 7
PAP2_SPOEX
ID PAP2_SPOEX STANDARD; PRT; 23 AA.
AC P30256;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PARALYTIC PEPTIDE II (PP II).
OS Spodoptera exigua (Beet armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
OX NCBI_TaxID=7107;
RN [1]
SEQUENCE.
RX MEDLINE=91302298; PubMed=2071576;
RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA Quistad G.B.;
RT "Isolation and identification of paralytic peptides from hemolymph of
RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
RT Heliothis virescens.";
RL J. Biol. Chem. 266:12873-12877(1991).
CC -1- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
CC HEMOLYPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC -1- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
DR PIR; D39855; E39855.
DR InterPro: IPR003463; GBP_PSP.
DR Pfam: PF02425; GBP_PSP; 1.
KW Hemolymph.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2505 MW; 0A96C85EB7D55AE0 CRC64;

Query Match 20.4%; Score 28.5; DB 1; Length 23;
Best Local Similarity 41.2%; Pred. No. 3.3e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 CGYGYQSIVD---RPDF 14
I III I :I I
Db 7 CTFGYQRTADGRCKPTF 23

RESULT 8
PAP3_SPOEX
ID PAP3_SPOEX STANDARD; PRT; 23 AA.
AC P30257;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PARALYTIC PEPTIDE III (PP III).
OS Spodoptera exigua (Beet armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
OX NCBI_TaxID=7107;
RN [1]
SEQUENCE.
RX MEDLINE=91302298; PubMed=2071576;
RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA Quistad G.B.;
RT "Isolation and identification of paralytic peptides from hemolymph of
RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
RT Heliothis virescens.";
RL J. Biol. Chem. 266:12873-12877(1991).
CC -1- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
CC HEMOLYPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC -1- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
DR PIR; E39855; E39855.
DR InterPro: IPR003463; GBP_PSP.
DR Pfam: PF02425; GBP_PSP; 1.
KW Hemolymph.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2505 MW; 0A96C85EB7D55AE0 CRC64;

Query Match 20.4%; Score 28.5; DB 1; Length 23;
Best Local Similarity 41.2%; Pred. No. 3.3e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 CGYGYQSIVD---RPDF 14
I III I :I I
Db 7 CTFGYQRTADGRCKPTF 23

RESULT 9
TP13_PHYRO
ID TP13_PHYRO STANDARD; PRT; 13 AA.
AC P04096;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 01-APR-1988 (Rel. 07, Last annotation update)
DE TRYPTOPHYLLIN-13.
OS Phyllomedusa rohdei (Rohde's leaf frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Phyllomedusa.
OX NCBI_TaxID=8394;
RN [1]
SEQUENCE.
RA Montecucchi P.C., Gozzini L., Erspamer V.;

```

```

FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2477 MW; 0A96CB4600855AE0 CRC64;

Query Match 20.4%; Score 28.5; DB 1; Length 23;
Best Local Similarity 41.2%; Pred. No. 3.3e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 CGYGYQSIVD---RPDF 14
I III I :I I
Db 7 CTFGYQRTADGRCKPTF 23

RESULT 8
PAP3_SPOEX
ID PAP3_SPOEX STANDARD; PRT; 23 AA.
AC P30257;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PARALYTIC PEPTIDE III (PP III).
OS Spodoptera exigua (Beet armyworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
OX NCBI_TaxID=7107;
RN [1]
SEQUENCE.
RX MEDLINE=91302298; PubMed=2071576;
RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
RA Quistad G.B.;
RT "Isolation and identification of paralytic peptides from hemolymph of
RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
RT Heliothis virescens.";
RL J. Biol. Chem. 266:12873-12877(1991).
CC -1- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
CC HEMOLYPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
CC -1- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
DR PIR; E39855; E39855.
DR InterPro: IPR003463; GBP_PSP.
DR Pfam: PF02425; GBP_PSP; 1.
KW Hemolymph.
FT DISULFID 7 19 BY SIMILARITY.
SQ SEQUENCE 23 AA; 2505 MW; 0A96C85EB7D55AE0 CRC64;

Query Match 20.4%; Score 28.5; DB 1; Length 23;
Best Local Similarity 41.2%; Pred. No. 3.3e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 3; Gaps 1;

QY 1 CGYGYQSIVD---RPDF 14
I III I :I I
Db 7 CTFGYQRTADGRCKPTF 23

RESULT 9
TP13_PHYRO
ID TP13_PHYRO STANDARD; PRT; 13 AA.
AC P04096;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-NOV-1986 (Rel. 03, Last sequence update)
DT 01-APR-1988 (Rel. 07, Last annotation update)
DE TRYPTOPHYLLIN-13.
OS Phyllomedusa rohdei (Rohde's leaf frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Phyllomedusa.
OX NCBI_TaxID=8394;
RN [1]
SEQUENCE.
RA Montecucchi P.C., Gozzini L., Erspamer V.;

```

RT "Primary structure determination of a tryptophan-containing
 RL tridecapeptide from Phyllomedusa rohdei";
 DR Int. J. Pept. Protein Res. 27:175-182(1986).
 KW PIR: A05174; A05174.

FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 SO SEQUENCE 13 AA; 1646 MW; 33BF33A212227773 CRC64;

Query Match 18.68; Score 26; DB 1; Length 13;
 Best Local Similarity 44.48; Pred. No. 4.3e+02;
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 10 DRDPFKPI 18
 : : : : :
 Db 2 EKPYWPPPI 10

RESULT 10
 PAPI_SPOEX
 ID PAPI_SPOEX STANDARD; PRT; 23 AA.
 AC P30255;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PARALYTIC PEPTIDE I (PP I).
 OS Spodoptera exigua (Beet armyworm).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Noctuoidea; Noctuidae; Amphipyrinae; Spodoptera.
 OX NCBI_TaxID=7107;
 RN [1]
 RC SEQUENCE.
 RC TISSUE=Hemolymph;
 RX MEDLINE=91302298; PubMed=2071576;
 RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
 RA Quistad G.B.;
 RT "Isolation and identification of paralytic peptides from hemolymph of
 RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
 RT Heliothis virescens.";
 RL J. Biol. Chem. 266:12873-12877(1991).
 CC -1- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
 CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
 CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
 CC -1- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
 DR PIR: C39855; C39855.
 DR InterPro: IPR003463; GBP_PSP.
 DR Pfam: PF02425; GBP_PSP; 1.
 KW Hemolymph.
 FT DISULFID 7 19 BY SIMILARITY.
 SQ SEQUENCE 23 AA; 2451 MW; 0A96D1F600855AE0 CRC64;

Query Match 18.68; Score 26; DB 1; Length 23;
 Best Local Similarity 50.08; Pred. No. 7.9e+02;
 Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGYGVQSIQD 10
 : : : : :
 Db 7 CTPGYQHTAD 16

RESULT 11
 CAMT_PINPS
 ID CAMT_PINPS STANDARD; PRT; 24 AA.
 AC P81081;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PROBABLE CAFFEYL-COA O-METHYLTRANSFERASE (EC 2.1.1.104) (TRANS-
 DE CAFFEYL-COA 3-O-METHYLTRANSFERASE) (CCOAMT) (WATER STRESS
 DE RESPONSIVE PROTEIN 13) (FRAGMENTS).
 OS Pinus pinaster (Maritime pine).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 OX NCBI_TaxID=71647;
 RN [1]
 RC SEQUENCE.
 RC TISSUE=Needle;
 RX MEDLINE=98418576; PubMed=9747804;
 RA Costa P., Bahrman N., Frigerio J.-M., Kremer A., Plomion C.;
 RA "Water-deficit-responsive proteins in maritime pine.";
 RL Plant Mol. Biol. 38:587-596(1998).
 RN [2]
 RC SEQUENCE.
 RC TISSUE=Needle;
 RX MEDLINE=99274088; PubMed=10344291;
 RA Costa P., Plomion C., Bauw G., Dubos C., Bahrman N., Kremer A.,
 RA Frigerio J.-M., Plomion C.;
 RA "Separation and characterization of needle and xylem maritime pine
 RT proteins.";
 RL Electrophoresis 20:1098-1108(1999).
 CC -1- FUNCTION: INVOLVED IN THE REINFORCEMENT OF THE PLANT CELL WALL
 CC UNDER CONDITIONS THAT TRIGGER THE DISEASE RESISTANCE RESPONSE (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + CAFFEYL-COA =
 CC S-ADENOSYL-L-HOMOCYSTEINE + FERULOYL-COA.
 CC -1- PATHWAY: LIGNIN BIOSYNTHESIS.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN
 CC IS: 5.2, ITS MW IS: 30 KDA.
 CC Lignin biosynthesis; transferase; Methyltransferase.
 KW NON_TER 1
 FT NON_TER 9 10
 FT NON_TER 24 24
 SQ SEQUENCE 24 AA; 2447 MW; AABE842F9EEB0CF0 CRC64;

Query Match 18.68; Score 26; DB 1; Length 24;
 Best Local Similarity 44.48; Pred. No. 8.3e+02;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 17 PIVRSITIC 25
 : : : : :
 Db 16 PVGGVTLIC 24

RESULT 12
 PAPI_MANSE
 ID PAPI_MANSE STANDARD; PRT; 23 AA.
 AC P30253;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PARALYTIC PEPTIDE I (PP I).
 OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Sphingioidea; Sphingidae; Sphinginae; Manduca.
 OX NCBI_TaxID=7130;
 RN [1]
 RC SEQUENCE.
 RC TISSUE=Hemolymph;
 RX MEDLINE=91302298; PubMed=2071576;
 RA Skinner W.S., Dennis P.A., Li J.P., Summerfelt R.M., Carney R.L.,
 RA Quistad G.B.;
 RT "Isolation and identification of paralytic peptides from hemolymph of
 RT the lepidopteran insects Manduca sexta, Spodoptera exigua, and
 RT Heliothis virescens.";
 RL J. Biol. Chem. 266:12873-12877(1991).
 CC -1- FUNCTION: CAUSES RAPID, RIGID PARALYSIS WHEN INJECTED INTO
 CC LEPIDOPTERAN LARVAE. THE PHYSIOLOGICAL ROLE MAY BE TO REDUCE
 CC HEMOLYMPH LOSS FOLLOWING INJURY AND PROMOTE WOUND HEALING.
 CC -1- SIMILARITY: BELONGS TO THE GBP / PSP1 / PARALYTIC PEPTIDE FAMILY.
 DR PIR: A39855; A39855.
 DR InterPro: IPR003463; GBP_PSP.

DR Pfam: PF02425; GBP_PSP; 1.
 KW Hemolymph.
 FT DISULFID 7 19 BY SIMILARITY.
 SQ SEQUENCE 23 AA; 2436 MW; 0B26CB5C29855FE4 CRC64;

Query Match 17.5%; Score 24.5; DB 1; Length 23;
 Best Local Similarity 35.3%; Pred. No. 1.3e+03;
 Matches 6; Conservative 1; Mismatches 7; Indels 3; Gaps 1;

QY 1 CGYGYQSVIVD---RPDF 14
 | | | | | | | | | |
 Db 7 CATGYLTADGRCKPTF 23

RESULT 13

CALL_CALGI STANDARD; PRT; 14 AA.
 AC P20728;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE CALOTROPIN DI (EC 3.4.22.-) (FRAGMENT).
 OS Calotropis gigantea (Madag) (Bowstring hemp).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Gentianales; Apocynaceae; Asclepiadoideae;
 OC Asclepiadeae; Calotropis.
 OX NCBI_TaxID=4066;
 RN [1]
 RP SEQUENCE.
 RA Bhattacharya D., Sengupta A., Sinha N.K.;
 RT "Chemical modification and amino terminal sequence of calotropin DI
 from Calotropis gigantea.";
 RL Phytochemistry 26:633-636(1987).
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1; ALSO KNOWN AS THE
 CC PAPAIN FAMILY OF THIOLE PROTEASES.
 DR PIR; PT0026; PT0026.
 DR MEROPS; C01.011; .
 DR InterPro; IPR000169; Thiolprot_act_site.
 DR PROSITE; PS00139; THIOLEPROTEASE_CYS; PARTIAL.
 DR PROSITE; PS00639; THIOLEPROTEASE_HIS; PARTIAL.
 DR PROSITE; PS00640; THIOLEPROTEASE_ASN; PARTIAL.
 KW Hydrolase; Thiol protease.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT NON_TER 14 14
 SQ SEQUENCE 14 AA; 1715 MW; D93F0276CDA4662 CRC64;

Query Match 17.1%; Score 24; DB 1; Length 14;
 Best Local Similarity 60.0%; Pred. No. 9.5e+02;
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 RPDPF 15
 | | | | |
 Db 2 RPEYP 6

RESULT 14

YMOR_PSEPU STANDARD; PRT; 16 AA.
 AC Q02210;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE HYPOTHETICAL PROTEIN IN MORA 3' REGION (FRAGMENT).
 OS Pseudomonas putida.
 OG Plasmid pMDH7.2.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas
 OX NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=M10;
 RX MEDLINE=93199531; PubMed=8452544;
 RA Willey D.L., Caswell D.A., Lowe C.R., Bruce N.C.;
 RT "Nucleotide sequence and over-expression of morphine dehydrogenase, a
 RT plasmid-encoded gene from Pseudomonas putida M10.";
 RL Biochem. J. 290:539-544(1993).
 CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; M94775; AAB17357.1; .
 DR PIR; S30384; S30384.
 KW Hypothetical protein; Plasmid.
 FT NON_TER 16 16
 SQ SEQUENCE 16 AA; 1962 MW; A2F1EB8C172766ED CRC64;

Query Match 17.1%; Score 24; DB 1; Length 16;
 Best Local Similarity 40.0%; Pred. No. 1.1e+03;
 Matches 6; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

QY 11 RDPFPPKIVRSITLC 25
 | | | | | | | | | |
 Db 3 RQDLPRS--RSLRSC 15

RESULT 15

UP21_UPEIN STANDARD; PRT; 19 AA.
 ID UP21_UPEIN
 AC P82027;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE UP21_UPEIN 2.1.
 DE UPERIN 2.1.
 OS Uperoleia inundata (Floodplain toadlet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
 OC Uperoleia.
 OX NCBI_TaxID=104953;
 RN [1]
 RP SEQUENCE, AND MASS SPECTROMETRY.
 RC Tissue=Skin secretion;
 RA Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
 RA Adams G.W., Severini C.;
 RT "Novel uperin peptides from the dorsal glands of the Australian
 RT floodplain toadlet Uperoleia inundata.";
 RL Aust. J. Chem. 49:475-484(1996).
 CC -I- FUNCTION: SHOWS A MEDIUM ANTIBACTERIAL ACTIVITY AGAINST
 CC L. MESENTERIOIDES, M. LUTEUS AND S. UBERIS.
 CC -I- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
 CC -I- MASS SPECTROMETRY: MW=1926; METHOD=FAB.
 KW Amphibian skin; Antibiotic.
 SQ SEQUENCE 19 AA; 1927 MW; 328834D77BA353D2 CRC64;

Query Match 17.1%; Score 24; DB 1; Length 19;
 Best Local Similarity 50.0%; Pred. No. 1.3e+03;
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 13 DFPPKIVRSI 22
 | | | | | | | | | |
 Db 4 DFAPKVVGGI 13

Search completed: March 4, 2002, 13:20:28
 Job time: 615 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:21:52 ; Search time 76.45 Seconds
(without alignments)
47.833 Million cell updates/sec

Title: US-09-701-623C-8
Perfect score: 140
Sequence: 1 GGYGYSIVDRPDPKPIVRSITLC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 7775

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL17.*
- 1: sp-archaea.*
 - 2: sp-bacteria.*
 - 3: sp-fungi.*
 - 4: sp-human.*
 - 5: sp-invertebrate.*
 - 6: sp-mammal.*
 - 7: sp-mhc.*
 - 8: sp-organelle.*
 - 9: sp-phage.*
 - 10: sp-plant.*
 - 11: sp-rodent.*
 - 12: sp-virus.*
 - 13: sp-vertebrate.*
 - 14: sp-unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 35 | 25.0 | 21 | 2 | Q9X3G4 prochloroco |
| 2 | 34 | 24.3 | 25 | 2 | O67969 prochloroco |
| 3 | 33 | 23.6 | 21 | 2 | Q9X3G0 prochloroco |
| 4 | 33 | 23.6 | 21 | 2 | Q9X3F2 prochloroco |
| 5 | 32 | 22.9 | 21 | 2 | Q9WV25 prochloroco |
| 6 | 32 | 22.9 | 21 | 9 | Q9MBH5 streptococ |
| 7 | 32 | 22.9 | 22 | 11 | Q9QV92 rattus sp. |
| 8 | 32 | 22.9 | 23 | 12 | O85496 bovine leuk |
| 9 | 31 | 22.1 | 14 | 10 | P82340 pisum sativ |
| 10 | 31 | 22.1 | 18 | 2 | O56610 vibrio chol |
| 11 | 31 | 22.1 | 21 | 2 | Q9X3H0 prochloroco |
| 12 | 30 | 21.4 | 21 | 2 | Q9X3J8 prochloroco |
| 13 | 29.5 | 21.1 | 24 | 5 | O9BM10 sponsilla l |
| 14 | 29 | 20.7 | 17 | 4 | O9UCE7 homo sapien |
| 15 | 28 | 20.0 | 20 | 12 | O41542 human immun |
| 16 | 28 | 20.0 | 21 | 2 | O9ZG55 chlamydia t |
| 17 | 28 | 20.0 | 21 | 2 | Q9X3K2 prochloroco |
| 18 | 28 | 20.0 | 21 | 2 | Q9X3G6 prochloroco |
| 19 | 28 | 20.0 | 21 | 2 | Q9X3E8 prochloroco |

| | | | | | |
|----|------|------|----|----|--------------------|
| 20 | 28 | 20.0 | 22 | 4 | O13659 homo sapien |
| 21 | 27 | 19.3 | 18 | 2 | Q9X3E9 prochloroco |
| 22 | 27 | 19.3 | 19 | 2 | O9R7I3 prochloroco |
| 23 | 27 | 19.3 | 20 | 2 | O67964 prochloroco |
| 24 | 27 | 19.3 | 21 | 2 | Q9X3K9 prochloroco |
| 25 | 27 | 19.3 | 21 | 2 | Q9X3I8 prochloroco |
| 26 | 27 | 19.3 | 21 | 2 | Q9X3I1 prochloroco |
| 27 | 27 | 19.3 | 21 | 2 | Q9X3H5 prochloroco |
| 28 | 27 | 19.3 | 21 | 2 | Q9X3D8 prochloroco |
| 29 | 27 | 19.3 | 21 | 2 | Q9WM38 prochloroco |
| 30 | 27 | 19.3 | 21 | 2 | Q9WM36 prochloroco |
| 31 | 27 | 19.3 | 21 | 2 | Q9WV23 prochloroco |
| 32 | 27 | 19.3 | 21 | 2 | Q9WVY2 prochloroco |
| 33 | 27 | 19.3 | 21 | 2 | Q9WWV5 prochloroco |
| 34 | 27 | 19.3 | 21 | 2 | Q9R2Y1 prochloroco |
| 35 | 27 | 19.3 | 21 | 4 | Q9BU87 homo sapien |
| 36 | 27 | 19.3 | 21 | 10 | O41559 triticum ae |
| 37 | 27 | 19.3 | 21 | 10 | O41564 triticum ae |
| 38 | 27 | 19.3 | 25 | 11 | O63993 rattus norv |
| 39 | 26.5 | 18.9 | 16 | 11 | O54894 mus musculu |
| 40 | 26.5 | 18.9 | 22 | 12 | O84254 bovine papi |
| 41 | 26.5 | 18.9 | 24 | 5 | O94374 caenorhabdi |
| 42 | 26 | 18.6 | 11 | 5 | O9NL65 ascaris suu |
| 43 | 26 | 18.6 | 17 | 10 | Q9S8S7 zea mays (m |
| 44 | 26 | 18.6 | 17 | 11 | Q9QV41 mus sp. spy |
| 45 | 26 | 18.6 | 18 | 12 | O41588 human immun |

ALIGNMENTS

RESULT 1

| | | | | |
|--------|-----------------------------------------------------------------------------------------------------------------------|--------------|----------|-------------------------|
| Q9X3G4 | O9X3G4 | PRELIMINARY; | PRT; | 21 AA. |
| AC | O9X3G4; | | | |
| DT | O1-NOV-1999 (Tremblrel. 12, Created) | | | |
| DT | O1-NOV-1999 (Tremblrel. 12, Last sequence update) | | | |
| DE | O1-NOV-1999 (Tremblrel. 12, Last annotation update) | | | |
| DE | CYTCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT) | | | |
| GN | PETD. | | | |
| OS | Prochlorococcus sp. | | | |
| OC | Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae; | | | |
| OC | Prochlorococcus. | | | |
| OX | NCBI_TaxID=1220; | | | |
| RN | {1} | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RA | Urbach E., Chisholm S.W.; | | | |
| RT | "Genetic diversity in Prochlorococcus populations flow cytometrically sorted from the Sargasso Sea and Gulf Stream."; | | | |
| RL | Limnol. Oceanog. 43:1615-1630(1998). | | | |
| DR | EMBL: AF070158; AAD23203.1; | | | |
| FT | NON_TER | | | |
| SQ | SEQUENCE | 21 AA; | 2340 MW; | CC947A36F12C854F CRC64; |

Query Match 25.0%; Score 35; DB 2; Length 21;
Best Local Similarity 46.7%; Pred. No. 1.4e+02;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

| | | | |
|----|---|----------------|----|
| QY | 7 | SIVDRPDPKPIVRS | 21 |
| | | : : | |
| Db | 2 | SILKKPDLTDPKLR | 16 |

RESULT 2

| | | | | |
|--------|-----------------------------------------------------|--------------|------|--------|
| O67969 | O67969 | PRELIMINARY; | PRT; | 25 AA. |
| ID | O67969 | | | |
| AC | O67969; | | | |
| DT | O1-AUG-1998 (Tremblrel. 07, Created) | | | |
| DT | O1-AUG-1998 (Tremblrel. 07, Last sequence update) | | | |
| DT | O1-NOV-1998 (Tremblrel. 08, Last annotation update) | | | |
| DE | B6/F COMPLEX SUBUNIT IV (FRAGMENT). | | | |
| GN | PETD. | | | |

OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MIT9303;
RX MEDLINE=98123172; PubMed=9452521;
RA Urbach E., Scanlan D.J., Distel D.L., Waterbury J.B., Chisholm S.W.;
RT "Rapid diversification of marine picophytoplankton with dissimilar
RT light-harvesting structures inferred from sequences of Prochlorococcus
RT and Synechococcus (Cyanobacteria).";
RL J. Mol. Evol. 46:188-201(1998).
DR EMBL; AF001491; AAC05629.1; -;
FT NON_TER 25
SQ SEQUENCE 25 AA; 2833 MW; 504CCD63AC947A21 CRC64;

Query Match 24.3%; Score 34; DB 2; Length 25;
Best Local Similarity 42.9%; Pred. No. 2.4e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 8 IVDPRDPFKPIVRS 21
I: :||| :|:
Db 3 ILKRPDSDPKLRA 16

RESULT 3
ID Q9X3G0 PRELIMINARY; PRT; 21 AA.
AC Q9X3G0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).
GN PETD.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanog. 43:1615-1630(1998).
DR EMBL; AF070155; AAD20780.1; -;
FT NON_TER 21
SQ SEQUENCE 21 AA; 2324 MW; D3AFDAECFA34E54F CRC64;

Query Match 23.6%; Score 33; DB 2; Length 21;
Best Local Similarity 33.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 7 SIVDRPDPFKPIVRS 21
I: :||| :|:
Db 2 SVIKKPDLSDPKRA 16

RESULT 4
ID Q9X3F2 PRELIMINARY; PRT; 21 AA.
AC Q9X3F2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).
GN PETD.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;

RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanog. 43:1615-1630(1998).
DR EMBL; AF070149; AAD20770.1; -;
FT NON_TER 21
SQ SEQUENCE 21 AA; 2338 MW; CC92DAFF3434E54F CRC64;

Query Match 23.6%; Score 33; DB 2; Length 21;
Best Local Similarity 33.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 7 SIVDRPDPFKPIVRS 21
I: :||| :|:
Db 2 SVIKKPDLEDPKRA 16

RESULT 5
ID Q9WVZ5 PRELIMINARY; PRT; 21 AA.
AC Q9WVZ5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).
GN PETD.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanog. 43:1615-1630(1998).
DR EMBL; AF070172; AAD20786.1; -;
DR EMBL; AF070170; AAD20783.1; -;
FT NON_TER 21
SQ SEQUENCE 21 AA; 2294 MW; CC947A26F034E54F CRC64;

Query Match 22.9%; Score 32; DB 2; Length 21;
Best Local Similarity 40.0%; Pred. No. 4e+02;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 7 SIVDRPDPFKPIVRS 21
I: :||| :|:
Db 2 SILKKPDLDPKLRA 16

RESULT 6
ID Q9MBH5 PRELIMINARY; PRT; 21 AA.
AC Q9MBH5; Q9MBY6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CELL WALL LYTIC ENZYME (FRAGMENT).
GN MML.
OS Streptococcus pneumoniae bacteriophage MML.
OC Viruses.
OX NCBI_TaxID=120574;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20392133; PubMed=10933687;
RA Gindreau E., Lopez R., Garcia P.;
RT "MML, a temperate bacteriophage of the type 23F Spanish/USA
RT multiresistant epidemic clone of Streptococcus pneumoniae: structural
RT analysis of the site-specific integration system.";

RL J. Virol. 74:7803-7813(2000).
DR EMBL; AJ400632; CAB96625.1; -.
DR EMBL; AJ400629; CAB96617.1; -.
FT NON_TER 1
SQ SEQUENCE 21 AA; 2235 MW; 30C8B7B7E0D8F244 CRC64;

Query Match 22.9%; Score 32; DB 9; Length 21;
Best Local Similarity 62.5%; Pred. No. 4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 SIVDRPDF 14
| : | : | : |
Db 3 SMADKPDF 10

RESULT 7

ID Q9QV92 PRELIMINARY; PRT; 22 AA.

AC Q9QV92;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE P93 CALCIUM-BINDING PROTEIN (FRAGMENT).

OS Rattus sp.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=101118;

RN [1]

RP SEQUENCE

RX MEDLINE=93179436; PubMed=8440713;

RA Gluchrist J.S., Pierce G.N.;

RT "Identification and purification of a calcium-binding protein in

RT hepatic nuclear membranes."

RL J. Biol. Chem. 268:4291-4299(1993).

DR InterPro; IPR001580; Calreticulin.

DR Pfam; PF00262; calreticulin; 1.

DR ProDom; PD001866; Calreticulin; 1.

SQ SEQUENCE 22 AA; 2407 MW; B82830E4477E25F9 CRC64;

Query Match 22.9%; Score 32; DB 11; Length 22;
Best Local Similarity 28.8%; Pred. No. 4.2e+02;
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGYQSIQIVDRPDF 14
| : | : | : |
Db 7 CGVQRPMDPNY 20

RESULT 8

ID Q85496 PRELIMINARY; PRT; 23 AA.

AC Q85496;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)

DE ENVELOPE PROTEIN (FRAGMENT).

OS Bovine leukemia virus (BLV).

OC Viruses; Retroid viruses; Retroviridae; Deltaretrovirus.

OX NCBI_TaxID=11901;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=87198886; PubMed=3033284;

RA Rice N.R., Simek S.L., Dubois G.C., Showalter S.D., Gilden R.V.,

RA Stephens R.W.;

RT "Expression of the bovine leukemia virus X region in virus-infected

RT cells."

RL J. Virol. 61:1577-1585(1987).

DR EMBL; M16017; AAA87336.1; -.

KW Envelope protein.

FT NON_TER 23

SQ SEQUENCE 23 AA; 2877 MW; F85E7CBED5440B08 CRC64;

Query Match 22.9%; Score 32; DB 12; Length 23;
Best Local Similarity 50.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 15 PKPIVRSITL 24
| : | : | : |
Db 11 PQPIRWVSL 20

RESULT 9

P82340

ID P82340 PRELIMINARY; PRT; 14 AA.

AC P82340;

DT 01-JUN-2000 (TrEMBLrel. 14, Created)

DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE UNKNOWN PROTEIN FROM 2D-PAGE OF THYLAKOID LUMEN (SPOT204)

DE (FRAGMENT).

OS Pisum sativum (Garden pea).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids I; Fabales; Fabaceae; Papilionoideae; Pisum.

OX NCBI_TaxID=3888;

RN [1]

RP SEQUENCE, SUBCELLULAR LOCATION, AND DEVELOPMENTAL STAGE.

RC STRAIN=CV. DE GRACE; TISSUE=LEAF;

RX MEDLINE=20181728; PubMed=10715320;

RA Peltier J.-B., Friso G., Kalume D.E., Roepstorff P., Nilsson F.,

RA Adamska I., van Wijk K.J.;

RT "Proteomics of the chloroplast: systematic identification and

RT targeting analysis of luminal and peripheral thylakoid proteins."

RL Plant Cell 12:319-341(2000).

CC -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE LUMEN.

CC -!- DEVELOPMENTAL STAGE: UNFOLDED AND FULLY DEVELOPED LEAVES.

KW Chloroplast; Thylakoid membrane.

FT NON_TER 14

SQ SEQUENCE 14 AA; 1505 MW; 2EABFAF980F3D7D7 CRC64;

Query Match 22.1%; Score 31; DB 10; Length 14;
Best Local Similarity 62.5%; Pred. No. 3.6e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 GYQSIQIVDR 11
| : | : | : |
Db 4 GQPVVDVR 11

RESULT 10

Q56610

ID Q56610 PRELIMINARY; PRT; 18 AA.

AC Q56610;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)

DE ACCA (FRAGMENT).

GN ACCA.

OS Vibrio cholerae.

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

OX NCBI_TaxID=666;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C6706;

RX MEDLINE=97074686; PubMed=8917113;

RA Franco A., Peir-En Y., Johnson J., Barry E.M., Guerra H., Maurer R.,

RA Morris J.G.;

RT "Cloning and characterization of dnaE, encoding the catalytic subunit

RT of replicative DNA polymerase III, from Vibrio cholerae strain

RL Gene 175:281-283(1996).

FT NON_TER 14

SQ SEQUENCE 14 AA; 1505 MW; 2EABFAF980F3D7D7 CRC64;

```
FT NON_TER 18
SQ SEQUENCE 18 AA; 2153 MW; 18EBCDAD212842EF CRC64;

Query Match      22.1%; Score 31; DB 2; Length 18;
Best Local Similarity 85.7%; Pred. No. 4.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 13 DPKPIV 19
   |||||
Db 7 DFERPIV 13

RESULT 11
Q9X3H0
ID Q9X3H0 PRELIMINARY; PRT; 21 AA.
AC Q9X3H0; 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).
GN PETD.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
   sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanogr. 43:1615-1630(1998).
DR EMBL: AF070163; AAD23212.1; -
FT NON_TER 21
SQ SEQUENCE 21 AA; 2280 MW; CC946C4A2034E54F CRC64;

Query Match      22.1%; Score 31; DB 2; Length 21;
Best Local Similarity 33.3%; Pred. No. 5.7e+02;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 7 SIVDRPDPFKPIVRS 21
   ||: ||| | : ||
Db 2 SVLKKPDLADPKLRA 16

RESULT 12
Q9X3J8
ID Q9X3J8 PRELIMINARY; PRT; 21 AA.
AC Q9X3J8; 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).
GN PETD.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
   sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanogr. 43:1615-1630(1998).
DR EMBL: AF070190; AAD23228.1; -
FT NON_TER 21
SQ SEQUENCE 21 AA; 2298 MW; CC946FFCC02C854F CRC64;

Query Match      21.4%; Score 30; DB 2; Length 21;
Best Local Similarity 40.0%; Pred. No. 8e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 7 SIVDRPDPFKPIVRS 21
   | : ||| | : ||
Db 2 STLKKPDLADPKLRS 16
```

```
RESULT 13
Q9BM10
ID Q9BM10 PRELIMINARY; PRT; 24 AA.
AC Q9BM10; 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE GYPSY-LIKE REVERSE TRANSCRIPTASE (FRAGMENT).
OS Spongilla lacustris (Freshwater sponge).
OC Eukaryota; Metazoa; Porifera; Demospongiae; Ceractinomorpha;
OC Haplosclerida; Spongillidae; Spongilla.
OX NCBI_TaxID=6055;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-GRT-G6 RETROTRANSPOSON;
RX MEDLINE=20570504; PubMed=11121049;
RA Arkhipova I., Meselson M.;
RT "Transposable elements in sexual and asexual taxa.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14473-14477(2000).
DR EMBL: AY013996; AAG59968.1; -
KW RNA-directed DNA polymerase.
FT NON_TER 1
   |
FT NON_TER 24
SQ SEQUENCE 24 AA; 2499 MW; 4A03DE8E3F7A4D12 CRC64;
```

Query Match 21.1%; Score 29.5; DB 5; Length 24;
Best Local Similarity 54.5%; Pred. No. 1.1e+03;
Matches 6; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

```
QY 2 GYGQSIIVDRP 12
   || : ||| : ||
Db 6 GY-FQEIMDQP 15
```

```
RESULT 14
Q9UCE7
ID Q9UCE7 PRELIMINARY; PRT; 17 AA.
AC Q9UCE7; 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE D(TTAGGG)N-BINDING PROTEIN B37-TYPE A-B HETEROGENEOUS NUCLEAR
   RIBONUCLEOPROTEIN HOMOLOG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=93309464; PubMed=8321232;
RA Ishikawa F., Matunis M.J., Dreyfuss G., Cech T.R.;
RT "Nuclear proteins that bind the pre-mRNA 3' splice site sequence
   r(UUAG/G) and the human telomeric DNA sequence d(TTAGGG)n.";
RL Mol. Cell. Biol. 13:4301-4310(1993).
SQ SEQUENCE 17 AA; 1830 MW; 2A4B0AC841A66C37 CRC64;
```

Query Match 20.7%; Score 29; DB 4; Length 17;
Best Local Similarity 38.5%; Pred. No. 9.1e+02;
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```
QY 5 YQSIIVDRPDPKP 17
   : ||| | | |
Db 1 FGEVVDXTDTPDP 13
```

```

RESULT 15
O41542
ID O41542 PRELIMINARY; PRT; 20 AA.
AC O41542;
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C10;
RX MEDLINE=98105804; PubMed=9445059;
RA Connor R.I., Korber B.T.M., Graham B.S., Hahn B.H., Ho D.D.,
RA Walker B.D., Neumann A.U., Vermund S.H., Mestecky J., Jackson S.,
RA Fenamore E., Cao Y., Gao F., Kalams S., Kunstman K.J., McDonald D.,
RA McWilliams N., Trkola A., Moore J.P., Wollinsky S.M.;
RT "Immunological and virological analyses of persons infected by human
RT recombinant gp120 subunit vaccines.";
RL J. Virol. 72:1552-1576(1998).
DR ENBL; U84802; AAC58830.1;...
DR InterPro: IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
KW Envelope protein.
FT NON_TER 1
SQ SEQUENCE 20 AA; 2208 MW; 656A8CB2AF83D46E CRC64;

Query Match 20.0%; Score 28; DB 12; Length 20;
Best Local Similarity 50.0%; Pred. No. 1.5e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 16 KPIVRSITLC 25
||| :||
Db 3 KPCVKLTSLC 12

```

Search completed: March 4, 2002, 13:21:53
Job time: 654 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:10:52 ; Search time 37.41 seconds
(without alignments)
15.038 Million cell updates/sec

Title: US-09-701-623C-8

Perfect score: 140

Sequence: 1 CGYQSVIVDRPDPFPIVRSITLC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 123821

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1 | 64 | 45.7 | 25 | 3 US-09-100-414B-95 | Sequence 95, Appl |
| 2 | 64 | 45.7 | 25 | 4 US-09-303-323-95 | Sequence 95, Appl |
| 3 | 38 | 27.1 | 16 | 1 US-08-447-010-14 | Sequence 14, Appl |
| 4 | 38 | 27.1 | 21 | 1 US-08-447-411-19 | Sequence 19, Appl |
| 5 | 38 | 27.1 | 21 | 1 US-08-447-411-59 | Sequence 59, Appl |
| 6 | 38 | 27.1 | 21 | 2 US-08-662-227-16 | Sequence 16, Appl |
| 7 | 38 | 27.1 | 21 | 4 US-09-017-947-16 | Sequence 16, Appl |
| 8 | 38 | 27.1 | 22 | 1 US-08-447-411-72 | Sequence 72, Appl |
| 9 | 38 | 27.1 | 22 | 2 US-08-662-227-30 | Sequence 30, Appl |
| 10 | 38 | 27.1 | 22 | 4 US-09-017-947-30 | Sequence 30, Appl |
| 11 | 36 | 25.7 | 20 | 4 US-08-602-999A-120 | Sequence 120, App |
| 12 | 35 | 25.0 | 21 | 1 US-08-447-411-58 | Sequence 58, Appl |
| 13 | 35 | 25.0 | 21 | 2 US-08-662-227-15 | Sequence 15, Appl |
| 14 | 35 | 25.0 | 21 | 4 US-09-017-947-15 | Sequence 15, Appl |
| 15 | 35 | 25.0 | 22 | 1 US-08-447-411-71 | Sequence 71, Appl |
| 16 | 35 | 25.0 | 22 | 2 US-08-146-028-101 | Sequence 101, App |
| 17 | 35 | 25.0 | 22 | 2 US-08-662-227-29 | Sequence 29, Appl |
| 18 | 35 | 25.0 | 22 | 4 US-08-723-425A-101 | Sequence 101, App |
| 19 | 35 | 25.0 | 22 | 4 US-09-112-206-101 | Sequence 101, App |
| 20 | 35 | 25.0 | 22 | 4 US-09-017-947-29 | Sequence 29, Appl |
| 21 | 32 | 22.9 | 13 | 6 517197-5 | Patent No. 517197 |
| 22 | 32 | 22.9 | 18 | 3 US-08-630-916A-109 | Sequence 109, App |
| 23 | 32 | 22.9 | 20 | 2 US-08-466-975A-18 | Sequence 18, Appl |
| 24 | 32 | 22.9 | 20 | 2 US-08-466-975A-19 | Sequence 19, Appl |
| 25 | 32 | 22.9 | 20 | 2 US-08-391-671A-18 | Sequence 18, Appl |
| 26 | 32 | 22.9 | 20 | 2 US-08-391-671A-19 | Sequence 19, Appl |
| 27 | 32 | 22.9 | 20 | 3 US-08-467-902A-18 | Sequence 18, Appl |

28 32 22.9 20 3 US-08-467-902A-19 Sequence 19, Appl
29 32 22.9 20 4 US-09-275-265-18 Sequence 18, Appl
30 32 22.9 20 4 US-09-275-265-19 Sequence 19, Appl
31 32 22.9 21 2 US-08-448-600-9 Sequence 9, Appl
32 32 22.9 21 3 US-08-937-610-17 Sequence 17, Appl
33 32 22.9 22 2 US-08-146-028-61 Sequence 61, Appl
34 32 22.9 22 2 US-08-146-028-62 Sequence 62, Appl
35 32 22.9 22 2 US-08-146-028-122 Sequence 122, App
36 32 22.9 22 4 US-08-723-425A-61 Sequence 61, Appl
37 32 22.9 22 4 US-08-723-425A-62 Sequence 62, Appl
38 32 22.9 22 4 US-08-723-425A-122 Sequence 122, App
39 32 22.9 22 4 US-09-112-206-61 Sequence 61, Appl
40 32 22.9 22 4 US-09-112-206-62 Sequence 62, Appl
41 32 22.9 22 4 US-09-112-206-122 Sequence 122, App
42 31 22.1 6 3 US-08-467-580-49 Sequence 49, Appl
43 31 22.1 9 2 US-08-146-028-421 Sequence 421, App
44 31 22.1 9 4 US-08-723-425A-421 Sequence 421, App
45 31 22.1 9 4 US-09-112-206-421 Sequence 421, App

ALIGNMENTS

RESULT 1
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-95

Query Match 45.7%; Score 64; DB 3; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0019;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYQSVIVDRPDPFPIVRSITLC 25
||| | | | | | | | | | | | | | | | | | | |

Db 1 CGYQSVIVDRPDPFPIVRSITLC 25
||| | | | | | | | | | | | | | | | | | | |

RESULT 2

US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 622897
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-303-323-95

Query Match 45.7%; Score 64; DB 4; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0019;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYQSIYDRDPKPKIVRSITLC 25
|| ||| | | | : ||| |
Db 1 CGETYQSRVTHPLPRALMRSTKTC 25

RESULT 3

US-08-447-010-14
; Sequence 14, Application US/08447010
; Patent No. 5770718
; GENERAL INFORMATION:
; APPLICANT: MOFFATT, BARBARA
; TITLE OF INVENTION: GENE FOR APRT FROM PLANT TISSUE
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, Suite 701
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,010
; FILING DATE: 22-MAY-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/230,695
; FILING DATE: 21-APR-1994
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,132
; FILING DATE: 26-MAY-1992
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, MICHAEL I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1811-183 MIS:V9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-447-010-14

Query Match 27.1%; Score 38; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PDFPKP 17
|||||
Db 3 PDFPKP 8

RESULT 4

US-08-447-411-19
; Sequence 19, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NO. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0

TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Naja naja
US-08-447-411-19

Query Match 27.1%; Score 38; DB 1; Length 21;
Best Local Similarity 42.9%; Pred. No. 15;
Matches 6; Conservative 4; Mismatches 4; Indels 4; Gaps 0;

QY 3 YGVQSVIVDRDPFK 16
:|:|:|:|:|:
Db 8 FGDDNIISRSDFPE 21

RESULT 5

US-08-447-411-59
Sequence 59, Application US/08447411
Patent No. 5773243

GENERAL INFORMATION:
APPLICANT: FRITZINGER, DAVID C.
APPLICANT: BREDEHORST, REINHARD
APPLICANT: VOGEL, CARL-WILHELM
TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,411
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/043,747
FILING DATE: 07-APR-1993

ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5773243man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-101-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Naja naja
US-08-447-411-59

Query Match 27.1%; Score 38; DB 1; Length 21;
Best Local Similarity 42.9%; Pred. No. 15;
Matches 6; Conservative 4; Mismatches 4; Indels 4; Gaps 0;

QY 3 YGVQSVIVDRDPFK 16
:|:|:|:|:|:
Db 8 FGDDNIISRSDFPE 21

RESULT 6

US-08-662-227-16
Sequence 16, Application US/08662227
Patent No. 5922320

GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
APPLICANT: BREDEHORST, REINHORST
APPLICANT: KOCK, MICHAEL
APPLICANT: FRITZINGER, DAVID
TITLE OF INVENTION: RECOMBINANT PROCVF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/662,227
FILING DATE: 14-JUN-1996
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-662-227-16

Query Match 27.1%; Score 38; DB 2; Length 21;
Best Local Similarity 42.9%; Pred. No. 15;
Matches 6; Conservative 4; Mismatches 4; Indels 4; Gaps 0;

QY 3 YGVQSVIVDRDPFK 16
:|:|:|:|:|:
Db 8 FGDDNIISRSDFPE 21

RESULT 7

US-09-017-947-16
Sequence 16, Application US/09017947
Patent No. 6303754

GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
APPLICANT: BREDEHORST, REINHORST
APPLICANT: KOCK, MICHAEL
TITLE OF INVENTION: RECOMBINANT PROCVF

NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/017,947
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/662,227
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptidic
US-09-017-947-16

Query Match 27.1%; Score 38; DB 4; Length 21;
Best Local Similarity 42.9%; Pred. No. 15;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YGQSVIVDRDPFK 16
Db 8 FGDDNIISRSDFPE 21

RESULT 8
US-08-447-411-72
Sequence 72, Application US/08447411
Patent No. 5773243
GENERAL INFORMATION:
APPLICANT: FRITZINGER, DAVID C.
APPLICANT: BREDEHORST, REINHARD
APPLICANT: VOGEL, CARL-WILHELM
TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
NUMBER OF SEQUENCES: 81
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,411
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/043,747
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NO. 5773243man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-101-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Naja naja
US-08-447-411-72

Query Match 27.1%; Score 38; DB 1; Length 22;
Best Local Similarity 42.9%; Pred. No. 15;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 3 YGQSVIVDRDPFK 16
Db 8 FGDDNIISRSDFPE 21

RESULT 9
US-08-662-227-30
Sequence 30, Application US/08662227
Patent No. 5922320
GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
APPLICANT: BREDEHORST, REINHARD
APPLICANT: KOCK, MICHAEL
TITLE OF INVENTION: FRITZINGER, DAVID
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/662,227
FILING DATE: 14-JUN-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: peptide
US-08-662-227-30

Query Match 27.1%; Score 38; DB 2; Length 22;
Best Local Similarity 42.9%; Pred. No. 15;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 YGQSVDRDPFK 16
:|:|:|:
Db 8 FGDDNIISRDPE 21

RESULT 10
US-09-017-947-30
; Sequence 30, Application US/09017947
; Patent No. 6303754
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHARDT
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/017,947
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/662,227
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-017-947-30

Query Match 27.1%; Score 38; DB 4; Length 22;
Best Local Similarity 42.9%; Pred. No. 15;
Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 3 YGQSVDRDPFK 16
:|:|:|:
Db 8 FGDDNIISRDPE 21

RESULT 11
US-08-602-999A-120
; Sequence 120, Application US/08602999A

Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWLKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; TITLE OF INVENTION: ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,999A
FILING DATE: 16-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-202
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 120:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-602-999A-120

Query Match 25.7%; Score 36; DB 4; Length 20;
Best Local Similarity 60.0%; Pred. No. 28;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 15 PKPIVRSITL 24
|||:|:|:
Db 9 PKPPIRSVSL 18

RESULT 12
US-08-447-411-58
; Sequence 58, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/447,411
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/043,747
;; FILING DATE: 07-APR-1993
;; ATTORNEY/AGENT INFORMATION:
;; NAME: OBLON, NO. 5773243man F.
;; REGISTRATION NUMBER: 24,618
;; REFERENCE/DOCKET NUMBER: 1126-101-0
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 413-3000
;; TELEFAX: (703) 413-2220
;; TELEX: 248855 OPAT UR
;; INFORMATION FOR SEQ ID NO: 58:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 21 amino acids
;; TYPE: aminc acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-447-411-58

Query Match 25.0%; Score 35; DB 1; Length 21;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 IVDPRDPFK 16
I: | | | | |
DB 13 IISRSDFPK 21

RESULT 13
US-08-662-227-15
; Sequence 15, Application US/08662227
; Patent No. 5922320
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/662,227
; FILING DATE: 14-JUN-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 15:

;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 21 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-08-662-227-15

Query Match 25.0%; Score 35; DB 2; Length 21;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 IVDPRDPFK 16
I: | | | | |
DB 13 IISRSDFPK 21

RESULT 14
US-09-017-947-15
; Sequence 15, Application US/09017947
; Patent No. 6303754
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,947
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/662,227
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-017-947-15

Query Match 25.0%; Score 35; DB 4; Length 21;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 IVDPRDPFK 16
I: | | | | |
DB 13 IISRSDFPK 21

RESULT 15
US-08-447-411-71
; Sequence 71, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-447-411-71

Query Match 25.0%; Score 35; DB 1; Length 22;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 8 IVDRDPFK 16
|: | | | |
Db 13 IISRDFFK 21

Search completed: March 4, 2002, 13:10:53
Job time: 300 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:11:40 ; Search time 40.99 Seconds
(without alignments)
46.459 Million cell updates/sec

Title: US-09-701-623c-84
Perfect score: 135
Sequence: 1 CGETYKSTVSHPDLPREVRSIAKC 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 4959

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 35 | 25.9 | 24 | 2 S29749 | serum albumin - do |
| 2 | 31 | 23.0 | 19 | 2 PS0186 | superoxide dismuta |
| 3 | 30 | 22.2 | 16 | 2 S30384 | hypothetical prote |
| 4 | 29 | 21.5 | 23 | 2 I28027 | superoxide dismuta |
| 5 | 29 | 21.5 | 25 | 2 JP0045 | ribosomal protein |
| 6 | 28 | 20.7 | 20 | 2 A61093 | glue protein - Cal |
| 7 | 27.5 | 20.4 | 15 | 2 PS0251 | 15K protein 5106 - |
| 8 | 27 | 20.0 | 11 | 2 S78765 | ribosomal protein |
| 9 | 27 | 20.0 | 23 | 2 S41390 | p7 protein - human |
| 10 | 26.5 | 19.6 | 25 | 2 A58647 | alphaA-conotoxin P |
| 11 | 26 | 19.3 | 9 | 2 I46016 | cytochrome 4 - bo |
| 12 | 26 | 19.3 | 15 | 4 I38335 | hypothetical TEL/M |
| 13 | 26 | 19.3 | 16 | 2 A49255 | T-cell receptor be |
| 14 | 26 | 19.3 | 19 | 2 PA0012 | superoxide dismuta |
| 15 | 26 | 19.3 | 23 | 2 A41263 | kinase-related tra |
| 16 | 26 | 19.3 | 24 | 2 S35641 | DNA (cytosine-5)- |
| 17 | 25 | 18.5 | 10 | 2 H37196 | bradykinin-potentl |
| 18 | 25 | 18.5 | 15 | 2 PA0071 | superoxide dismuta |
| 19 | 25 | 18.5 | 17 | 2 PH1822 | T cell receptor al |
| 20 | 25 | 18.5 | 20 | 2 D49164 | chromogranin-B - r |
| 21 | 25 | 18.5 | 21 | 2 S47207 | T-cell receptor J- |
| 22 | 25 | 18.5 | 23 | 2 F61491 | seed protein ws-6 |
| 23 | 24.5 | 18.1 | 24 | 2 S42785 | relaxin - baboon (|
| 24 | 24 | 17.8 | 10 | 2 B37196 | bradykinin-potentl |
| 25 | 24 | 17.8 | 14 | 2 S48685 | extension protein |
| 26 | 24 | 17.8 | 14 | 2 PL0152 | metal-binding prot |
| 27 | 24 | 17.8 | 18 | 2 PC2280 | prolylendopeptidas |
| 28 | 24 | 17.8 | 20 | 2 S23981 | outer layer protei |
| 29 | 24 | 17.8 | 20 | 2 A61276 | superoxide dismuta |

30 24 17.8 23 2 S23637 hypothetical prote
31 24 17.8 25 2 JP0064 ribosomal protein
32 23 17.0 13 2 A54326 glandular kallikre
33 23 17.0 14 1 NTKNIM alpha-conotoxin MI
34 23 17.0 14 2 S74128 superoxide dismuta
35 23 17.0 14 2 PH1806 T cell receptor al
36 23 17.0 15 2 PA0009 seed storage prote
37 23 17.0 15 2 PA0088 protein QF200051 -
38 23 17.0 17 2 PH1802 T cell receptor al
39 23 17.0 18 2 S18386 110K protein - mou
40 23 17.0 19 2 PQ0409 RNA-directed RNA p
41 23 17.0 20 2 S28435 major outer membra
42 23 17.0 20 2 T01691 hypothetical prote
43 23 17.0 20 2 A05310 apolipoprotein E -
44 23 17.0 21 2 A56901 nerve growth facto
45 23 17.0 23 2 S72535 probable acr-2 reg

ALIGNMENTS

RESULT 1
S29749
serum albumin - dog
C:Species: Canis lupus familiaris (dog)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999
C:Accession: S29749
R: Dixon, J.W.; Sarkar, B.
J. Biol. Chem. 249, 5872-5877, 1974
A:Title: Isolation, amino acid sequence and copper(II)-binding properties of peptide
A:Reference number: S29749; MUID:75011422
A:Accession: S29749
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-24 <DIX>
C:Superfamily: serum albumin; serum albumin repeat homology

Query Match 25.9%; Score 35; DB 2; Length 24;
Best Local Similarity 40.9%; Pred. No. 93;
Matches 9; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

Qy 3 ETYKSTVSH--PDLPREVRSI 22
| | | | : | | | | :
Db 1 EAYKSEIAHRYNDLGEHFRGL 22

RESULT 2
PS0186
superoxide dismutase (EC 1.15.1.1) (Mn) - rice (fragment)
C:Species: Oryza sativa (rice)
C:Date: 05-Mar-1993 #sequence_revision 05-Mar-1993 #text_change 05-Mar-1999
C:Accession: PS0186
R: Kamo, M.; Tsugita, A.
submitted to JIPID, June 1991
A:Reference number: PS0184
A:Accession: PS0186
A:Molecule type: protein
A:Residues: 1-19 <KAM>
A:Experimental source: callus
C:Function:
C:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxyg
C:Superfamily: superoxide dismutase (Mn)
C:Keywords: manganese; metalloprotein; oxidoreductase

Query Match 23.0%; Score 31; DB 2; Length 19;
Best Local Similarity 75.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 8 TVSHPDLP 15
| | : | | | |
Db 3 TVALPDLP 10

C:Accession: S78765
R:Graack, H.R.

Submitted to the Protein Sequence Database, July 1999

A:Reference number: S78760

A:Accession: S78765

A:Molecule type: protein

A:Residues: 1-11 <GRA>

C:Keywords: mitochondrion

F:1-11/Product: ribosomal protein MRP-S24 (fragment) #status experimental <MAT>

Query Match 20.0%; Score 27; DB 2; Length 11;
Best Local Similarity 30.0%; Pred. No. 6.4e+02;
Matches 3; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 11 HPDLPREVVR 20

DB 2 HVDVPKDLTK 11

RESULT 9

S41390

p7 protein - human adenovirus 3 (fragment)

C:Species: Mastadenovirus h3 (human adenovirus 3)

C:Date: 06-Jan-1995 #sequence_revision 30-Jan-1998 #text_change 08-Oct-1999

C:Accession: S41390

R:Cuzange, A.; Chroboczek, J.; Jacrot, B.

Submitted to the EMBL Data Library, January 1994

A:Description: The penton base of human adenovirus type 3 has the RGD motif.

A:Reference number: S41388

A:Accession: S41390

A:Molecule type: DNA

A:Residues: 1-23 <GUZ>

A:Cross-references: EMBL:229487; NID:g444048; PIDN:CAA82623.1; PID:g444051

A:Experimental source: serotype 3

Query Match 20.0%; Score 27; DB 2; Length 23;
Best Local Similarity 41.7%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGETYKSTVSH 12

DB 6 CMEAPASSTQHP 17

RESULT 10

A58647

alpha-conotoxin PIVA [validated] - cone shell (Conus purpurascens)

C:Species: Conus purpurascens (purple cone)

C:Date: 31-Oct-1997 #sequence_revision 07-Nov-1997 #text_change 15-Sep-2000

C:Accession: A58647

R:Hopkins, C.; Grille, M.; Miller, C.; Shon, K.J.; Cruz, L.J.; Gray, W.R.; Dykert, J.;

J. Biol. Chem. 270, 22361-22367, 1995

A:Title: A new family of Conus peptides targeted to the nicotinic acetylcholine receptor

A:Reference number: A58647; MUID:95403432

A:Accession: A58647

A:Molecule type: protein

A:Residues: 1-25 <HOP>

R:Han, K.H.; Hwang, K.J.; Kim, S.M.; Kim, S.K.; Gray, W.R.; Olivera, B.M.; Rivier, J.; S

submitted to the Brookhaven Protein Data Bank, December 1996

A:Reference number: A67666; PDB:1p1p

A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue

R:Han, K.H.; Hwang, K.J.; Kim, S.M.; Kim, S.K.; Gray, W.R.; Olivera, B.M.; Rivier, J.; S

Biochemistry 36, 1669-1677, 1997

A:Title: NMR structure determination of a novel conotoxin, [Pro 7,13] alpha A-conotoxin

A:Reference number: A58646; MUID:97200721

A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR

C:Superfamily: unassigned conotoxins

C:Keywords: acetylcholine receptor inhibitor; amidated carboxyl end; hydroxyproline; pos

F:2-16, 3-11, 14-23/Disulfide bonds: #status experimental

F:7,13/Modified site: 4-hydroxyproline (Pro) (partial) #status experimental

F:20/Modified site: 4-hydroxyproline (Pro) #status experimental

F:25/Modified site: amidated carboxyl end (Gln) #status experimental

Query Match 19.6%; Score 26.5; DB 2; Length 25;
Best Local Similarity 41.7%; Pred. No. 1.8e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1 CGETYKSTVSH 12

DB 3 CG-SYPNAACHP 13

RESULT 11

I46016

cytokerin 4 - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 05-Nov-1999

C:Accession: I46016

R:Blessing, M.; Jorcano, J.L.; Franke, W.W.

EMBO J. 8, 117-126, 1989

A:Title: Enhancer elements directing cell-type-specific expression of cyokeratin gen

A:Reference number: I46016; MUID:89231609

A:Accession: I46016

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-9 <BLE>

A:Cross-references: EMBL:X14478; NID:g303; PIDN:CAA32640.1; PID:g577897

Query Match 19.3%; Score 26; DB 2; Length 9;
Best Local Similarity 83.3%; Pred. No. 2.2e+05;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 TYKSTV 9

DB 2 SYKSTV 7

RESULT 12

I38335

hypothetical TEL/MNL mutant fusion protein type II - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 20-Apr-2000

C:Accession: I38335

R:Buijs, A.; Sherr, S.; van Baal, S.; van Bezouw, S.; van der Plas, D.; Van Kessel, A

Oncogene 10, 1511-1519, 1995

A:Title: Translocation (12;22) (p13;q11) in myeloproliferative disorders results in f

A:Reference number: I38031; MUID:95249265

A:Accession: I38335

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-15 <BUI>

A:Cross-references: EMBL:X85024; NID:g971471; PIDN:CAA59397.1; PID:g971472

C:Comment: This sequence is the chimeric product of a translocation mutation.

C:Genetics:

A:Gene: ETV6/MNL; TEL/MNL

A:Map position: 22q11/12p13

C:Keywords: fusion protein

Query Match 19.3%; Score 26; DB 4; Length 15;
Best Local Similarity 54.5%; Pred. No. 1.3e+03;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 5 YKSTVSHPDLP 15

DB 2 YRSPHSAHDL 12

RESULT 13

A49255

T-cell receptor beta chain V-D-J-C region (V beta 7, J beta 1.6) - human (fragment)

C:Species: Homo sapiens (man)

C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997
C;Accession: A49255
R;Rosenberg, W.M.; Moss, P.A.; Bell, J.I.
Eur. J. Immunol. 22, 541-549, 1992
A;Title: Variation in human T cell receptor V beta and J beta repertoire: analysis using
A;Reference number: A49039; MUID:92164737
A;Accession: A49255
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-16 <ROS>
A;Note: sequence extracted from NCBI backbone (NCBIP:90722)
C;Keywords: T-cell receptor

Db 4 RSSCEDPGCPRDEERA 19

Search completed: March 4, 2002, 13:11:40
Job time: 292 sec

Query Match 19.3%; Score 26; DB 2; Length 16;
Best Local Similarity 33.3%; Pred. No. 1.3e+03;
Matches 4; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHP 12
| : | | | |
Db 1 CASSYPGTQNSP 12

RESULT 14

PA0012
superoxide dismutase (EC 1.15.1.1) (Mn) - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Jun-1992; #sequence_revision 06-Jan-1995 #text_change 03-May-1996
C;Accession: PA0012
R;Kano, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.
submitted to JIPID, July 1994
A;Description: Separation and characterization of Arabidopsis proteins by two-dimensional
A;Reference number: PA0001
A;Accession: PA0012
A;Molecule type: protein
A;Residues: 1-19 <XAM>
A;Experimental source: callus
C;Superfamily: superoxide dismutase (Mn)
C;Keywords: manganese; metalloprotein; oxidoreductase

Query Match 19.3%; Score 26; DB 2; Length 19;
Best Local Similarity 62.5%; Pred. No. 1.6e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 TVSHPDLF 15
| : | | | |
Db 3 TFLPDLF 10

RESULT 15

A41263
kinase-related transforming protein (hck) (EC 2.7.1.-) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 28-May-1992 #sequence_revision 28-May-1992 #text_change 21-Mar-1996
C;Accession: A41263
R;Lock, P.; Ralph, S.; Stanley, E.; Boulet, I.; Ramsay, R.; Dunn, A.R.
Mol. Cell. Biol. 11, 4363-4370, 1991
A;Title: Two isoforms of murine hck, generated by utilization of alternative translation
A;Reference number: A41263; MUID:91342636
A;Accession: A41253
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-23 <LOC>
C;Keywords: phosphotransferase

Query Match 19.3%; Score 26; DB 2; Length 23;
Best Local Similarity 31.2%; Pred. No. 2e+03;
Matches 5; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 6 KSTVSHFDLPREVRS 21
: | | | | | | |

THIS PAGE BLANK (USPTO,

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:20:28 ; Search time 24.21 Seconds
(without alignments)
37.861 Million cell updates/sec

Title: US-09-701-623c-84

Perfect score: 135

Sequence: 1 GGETYKSTVSHPDLPREVRSIAKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 1446

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SWISSPROT_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 32 | 23.7 | 20 | 1 CS21_STRTR | P81621 streptococ |
| 2 | 30 | 22.2 | 16 | 1 YMOR_PSEPU | Q02210 pseudomonas |
| 3 | 27 | 20.0 | 20 | 1 SODM_HORVU | P28524 hordeum vul |
| 4 | 26.5 | 19.6 | 25 | 1 CXA4_CONPU | P55963 conus purpu |
| 5 | 26 | 19.3 | 21 | 1 CFPA_TREPH | P56738 treponema p |
| 6 | 25 | 18.5 | 10 | 1 BPP8_BOTIN | P30426 bothrops in |
| 7 | 25 | 18.5 | 13 | 1 SODM_CANFA | P54712 canis famil |
| 8 | 25 | 18.5 | 14 | 1 SODM_STRGR | P80732 streptomyce |
| 9 | 25 | 18.5 | 17 | 1 EFG_THEAO | Q01697 thermus aqu |
| 10 | 25 | 18.5 | 20 | 1 UN05_PINPS | P81674 pinus pinas |
| 11 | 24 | 17.8 | 10 | 1 BPP2_BOTIN | P30422 bothrops in |
| 12 | 24 | 17.8 | 20 | 1 VMO2_CHICK | Q9ps49 gallus gall |
| 13 | 24 | 17.8 | 23 | 1 SODM_RANCA | P36215 rana catesb |
| 14 | 23 | 17.0 | 14 | 1 CXAL_CONNA | P01521 conus magus |
| 15 | 23 | 17.0 | 15 | 1 MCA2_RHOOP | P56870 rhodococcus |
| 16 | 23 | 17.0 | 24 | 1 CXOA_CONST | P28880 conus stria |
| 17 | 23 | 17.0 | 24 | 1 LEC_CROJU | P16352 crotalaria |
| 18 | 22 | 16.3 | 10 | 1 ODP2_BOVIN | P11180 bos taurus |
| 19 | 22 | 16.3 | 13 | 1 BLAC_STRGR | P81173 streptomyce |
| 20 | 22 | 16.3 | 13 | 1 EP65_HUMAN | P54963 homo sapien |
| 21 | 22 | 16.3 | 15 | 1 VORA_METTM | P80907 methanobact |
| 22 | 22 | 16.3 | 16 | 1 ODO2_BOVIN | P11179 bos taurus |
| 23 | 22 | 16.3 | 19 | 1 UP22_UPEIN | P82028 uperoleia i |
| 24 | 22 | 16.3 | 19 | 1 UP23_UPEIN | P82029 uperoleia i |
| 25 | 22 | 16.3 | 20 | 1 HELT_HELHO | P46693 heloderma h |
| 26 | 22 | 16.3 | 20 | 1 MIL7_BOVIN | P35451 bos taurus |
| 27 | 22 | 16.3 | 21 | 1 GYRA_STRSH | P50071 streptomyce |
| 28 | 22 | 16.3 | 23 | 1 UHA4_HUMAN | P49289 homo sapien |
| 29 | 22 | 16.3 | 25 | 1 ANDT_ANDAU | P56684 androctonus |
| 30 | 21 | 15.6 | 11 | 1 TKNL_PSEGU | P42986 pseudophryn |
| 31 | 21 | 15.6 | 14 | 1 CXAL_CONCN | P56973 conus conso |
| 32 | 21 | 15.6 | 16 | 1 LPHT_ECOLI | P3058 escherichia |
| 33 | 21 | 15.6 | 16 | 1 MMPX_SOLTI | P80501 solanum tub |

34 21 15.6 17 1 UP34_UPEMJ
35 21 15.6 17 1 UP37_UPEMJ
36 21 15.6 19 1 CXA2_CONST
37 21 15.6 20 1 COXM_THUOB
38 21 15.6 20 1 COXM_THUOB
39 21 15.6 22 1 CYTB_THETS
40 21 15.6 23 1 XCL1_ACIGB
41 21 15.6 25 1 FRHB_METBA
42 20 14.8 13 1 CXAL_CONST
43 20 14.8 13 1 ECDE_LYMDI
44 20 14.8 15 1 CXAL_CONGE
45 20 14.8 18 1 AZM_OCTVU

ALIGNMENTS

RESULT 1
CS21_STRTR STANDARD; PRT; 20 AA.
AC P81621;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 21 KDA COLD-SHOCK INDUCED PROTEIN (FRAGMENT).
OS Streptococcus thermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1308;
RN [1]
RP SEQUENCE.
RC STRAIN=PB18;
RX MEDLINE=99456673; PubMed=10525839;
RA Perrin C., Guilmont C., Bracqart P., Gaillard J.L.;
RT "Expression of a new cold shock protein of 21.5 kDa and of the major
RL cold shock protein by Streptococcus thermophilus after cold shock.";
RL Curr. Microbiol. 39:342-0347(1999).
CC -!- INDUCTION: BY COLD SHOCK.
FT NON_TER 20
SQ SEQUENCE 20 AA; 2390 MW; 9C2CA57F266B80AD CRC64;

Query Match 23.7%; Score 32; DB 1; Length 20;
Best Local Similarity 66.7%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 ETYKSTVSH 11
|||
Db 6 ETIKETVNH 14

RESULT 2
YMOR_PSEPU STANDARD; PRT; 16 AA.
AC Q02210;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE HYPOTHETICAL PROTEIN IN MORA 3' REGION (FRAGMENT).
OS Pseudomonas putida.
OG Plasmid pMDH7.2.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M10;
RX MEDLINE=93199531; PubMed=8452544;
RA Willey D.L., Caswell D.A., Lowe C.R.;
RT "Nucleotide sequence and over-expression of morphine dehydrogenase, a
RT plasmid-encoded gene from Pseudomonas putida M10.";
RL Biochem. J. 290:539-544(1993).
CC -----


```
CC -!- FUNCTION: COMPONENT OF THE CYTOPLASMIC FILAMENTS THAT RUN THE
CC LENGTH OF THE ORGANISM JUST UNDERNEATH THE CYTOPLASMIC MEMBRANE.
CC -!- SUBCELLULAR LOCATION: AN ARRAY OF 4 TO 6 FILAMENTS LIE IN CLOSE
CC APPosition TO THE INNER MEMBRANE AND ARE ALWAYS LOCALIZED DIRECTLY
CC UNDERNEATH THE CORRESPONDING GROUP OF PERIPLASMIC FLAGELLA.
KW Structural protein; Antigen.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2231 MW; 574604B4FFC2D017 CRC64;

Query Match 19.3%; Score 26; DB 1; Length 21;
Best Local Similarity 50.0%; Pred. No. 6.3e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 VSHPDLPREV 18
   | | | | |
DB 11 VFHPKPSAV 20

RESULT 6
BPP8_BOTIN STANDARD; PRT; 10 AA.
AC P30426;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE BRADYKININ-POTENTIATING PEPTIDE S5,1 (ANGIOTENSIN-CONVERTING
DE ENZYME INHIBITOR).
OS Bothrops insularis (Island jararaca) (Queimada jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=90351557; PubMed=2386615;
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
RT "Primary structure and biological activity of bradykinin potentiating
RT peptides from Bothrops insularis snake venom.";
RL J. Protein Chem. 9:221-227(1990).
CC -!- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
DR PIR; H37196; H37196.
KW Hypotensive agent; Venom.
FT MOD_RES 1 1 PYBROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 10 AA; 1173 MW; 2FF835545761F6D8 CRC64;

Query Match 18.5%; Score 25; DB 1; Length 10;
Best Local Similarity 60.0%; Pred. No. 3.9e+02;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 11 HPDLP 15
   | | | |
DB 5 HPNIP 9

RESULT 7
SODM_CANFA STANDARD; PRT; 13 AA.
ID SODM_CANFA
AC P54712;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE SUPEROXIDE DISMUTASE [MN], MITOCHONDRIAL (EC 1.15.1.1) (FRAGMENT).
GN SOD2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
```

```
RN SEQUENCE.
RP TISSUE=Heart;
RX MEDLINE=98163340; PubMed=9504812;
RA Dunn M.J., Corbett J.M., Wheeler C.H.;
RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT dog heart proteins.";
RL Electrophoresis 18:2795-2802(1997).
CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
DR HSP; P04179; IMSD.
DR HSC-2DPAGE; P54712; DOG.
DR InterPro; IPR001189; SOD_MI.
DR Pfam; PF00081; sodfr; 1.
DR PROSITE; PS00088; SOD_MN; PARTIAL.
KW Oxidoreductase; Manganese; Mitochondrion.
FT NON_TER 13
SQ SEQUENCE 13 AA; 1476 MW; 9C9651DE8BE0672A CRC64;

Query Match 18.5%; Score 25; DB 1; Length 13;
Best Local Similarity 83.3%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 SHPDLP 15
   | | | | |
DB 3 SLPDLP 8

RESULT 8
SODN_STRGR STANDARD; PRT; 14 AA.
ID SODN_STRGR
AC P80732;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SUPEROXIDE DISMUTASE [NI] (EC 1.15.1.1) (NISOD) (NICKEL-CONTAINING
DE SUPEROXIDE DISMUTASE) (FRAGMENT).
GN SODN OR SODI.
OS Streptomyces griseus.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1911;
RN [1]
RP SEQUENCE.
RC STRAIN=KCTC 9006;
RX MEDLINE=97056064; PubMed=8900409;
RA Youn H.-D., Youn H., Lee J.-W., Yim Y.-I., Lee J.K., Hah Y.C.,
RA Kang S.-O.;
RT "Unique isozymes of superoxide dismutase in Streptomyces griseus.";
RL Arch. Biochem. Biophys. 334:341-348(1996).
CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: NICKEL.
CC -!- SUBUNIT: HOMOTETRAMER (POTENTIAL).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
KW Oxidoreductase; Nickel.
FT NON_TER 14
SQ SEQUENCE 14 AA; 1457 MW; 305C93EF783F2AC8 CRC64;

Query Match 18.5%; Score 25; DB 1; Length 14;
Best Local Similarity 62.5%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 11 HPDLPREV 18
   | | | | |
DB 1 HSDDLPGCV 8
```

```
RESULT 9
EFG_THEAQ
ID EFG_THEAQ STANDARD; PRT; 17 AA.
AC Q01697;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ELONGATION FACTOR G (EF-G) (FRAGMENT).
GN FUSA OR FUS.
OS Thermus aquaticus.
OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
OX NCBI_TaxID=271;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EP 00276;
RA MEDLINE=92362620; PubMed=1499561;
RA Voss R.H., Hartmann R.K., Lippmann C., Alexander C., Jahn O.,
RA Erdmann V.;
RT "Sequence of the tufA gene encoding elongation factor EF-Tu from
RT Thermus aquaticus and overproduction of the protein in Escherichia
RT coli.";
RL Eur. J. Biochem. 207:839-846(1992).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
CC RIBOSOME.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X66322; JAA46997.1; -
DR PIR; S29294; S29294.
DR HSP; P13551; IELO.
DR InterPro; IPR000795; GTP_EFTU.
DR PROSITE; PS00301; EFATOR_GTP; PARTIAL.
KW Elongation factor; Protein biosynthesis; GTP-binding.
FT NON_TER 1 1
SQ SEQUENCE 17 AA; 2094 MW; EA46E1EF05F86E1D CRC64;

Query Match 18.5%; Score 25; DB 1; Length 17;
Best Local Similarity 21.4%; Pred. No. 7e+02;
Matches 3; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 11 HPDLPREVRSIAK 24
: : : : :
DB 2 YQEVPRQICEKLIK 15

RESULT 10
UN05_PINPS
ID UN05_PINPS STANDARD; PRT; 20 AA.
AC P81674;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF NEEDLES (N147) (FRAGMENTS).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
OX NCBI_TaxID=71647;
RN [1]
RP SEQUENCE.
RC TISSUP=Needle.
RX MEDLINE=99274088; PubMed=10344291;
```

```
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
RT proteins.";
RL Electrophoresis 20:1098-1108(1999).
CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.6, ITS MW IS: 36 KDA.
FT NON_TER 1 1
FT NON_CONS 11 12 F -> I.
FT VARIANT 13 13 Y -> E.
FT VARIANT 14 14 R -> K.
FT VARIANT 15 15
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 2438 MW; 9F4E4678E086C298 CRC64;

Query Match 18.5%; Score 25; DB 1; Length 20;
Best Local Similarity 80.0%; Pred. No. 8.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 PDLPR 16
: : : :
DB 16 PELPR 20

RESULT 11
BPP2_BOTIN
ID BPP2_BOTIN STANDARD; PRT; 10 AA.
AC P30422;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE BRADYKININ-POTENTIATING PEPTIDE S4,3,1 (10C) (ANGIOTENSIN-CONVERTING
DE ENZYME INHIBITOR).
OS Bothrops insularis (Island jararaca) (Queimada jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=90351557; PubMed=2386615;
RA Cintra A.C.O., Vieira C.A., Giglio J.R.;
RT "Primary structure and biological activity of bradykinin potentiating
RT peptides from Bothrops insularis snake venom.";
RL J. Protein Chem. 9:221-227(1990).
CC -1- FUNCTION: THIS PEPTIDE BOTH INHIBITS THE ACTIVITY OF THE
CC ANGIOTENSIN-CONVERTING ENZYME AND ENHANCES THE ACTION OF
CC BRADYKININ BY INHIBITING THE KINASES THAT INACTIVATE IT.
CC IT ACTS AS AN INDIRECT HYPOTENSIVE AGENT.
DR PIR; B37196; B37196.
KW Hypotensive agent; Venom.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
SQ SEQUENCE 10 AA; 1213 MW; 30C53546C761F773 CRC64;

Query Match 17.8%; Score 24; DB 1; Length 10;
Best Local Similarity 60.0%; Pred. No. 5.5e+02;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 HPDLP 15
: : : :
DB 5 HPQIP 9

RESULT 12
VMO2_CHICK
ID VMO2_CHICK STANDARD; PRT; 20 AA.
AC Q9PS49;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
```

DE VITELLINE MEMBRANE OUTER LAYER PROTEIN II (VMO-II) (VMOII) (FRAGMENT).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Egg yolk;
 RX MEDLINE=92392273; PubMed=1520265;
 RA Kido S., Morimoto A., Kim F., Doi Y.;
 RT "Isolation of a novel protein from the outer layer of the vitelline
 membrane.";
 RL Biochem. J. 286:17-22(1992).
 CC -!- FUNCTION: EXACT FUNCTION NOT KNOWN, COMPONENT OF THE OUTER
 CC MEMBRANE OF THE VITELLINE LAYER OF THE EGG.
 CC -!- PTM: ALL CYSTEINE RESIDUES OF THE MATURE PROTEIN ARE INVOLVED IN
 CC DISULFIDE BONDS.
 CC NON_TER 20
 FT SEQUENCE 20 AA; 2325 MW; 45FC7989AB7527C7 CRC64;
 SQ

Query Match 17.8%; Score 24; DB 1; Length 20;
 Best Local Similarity 44.4%; Pred. No. 1.2e+03;
 Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 14 LPREVRSI 22
 |||: |:
 Db 1 LPRDTSRXV 9

RESULT 13
 ID SODM_RANCA STANDARD; PRT; 23 AA.
 AC P36215;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DE 20-AUG-2001 (Rel. 40, Last annotation update)
 DE SUPEROXIDE DISMUTASE [MN], MITOCHONDRIAL (EC 1.15.1.1) (FRAGMENT).
 OS Rana catesbeiana (Bull frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
 OX NCBI_TaxID=8400;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=87126854; PubMed=3492965;
 RA Abe Y., Okazaki T.;
 RT "Purification and properties of the manganese superoxide dismutase
 from the liver of bullfrog, Rana catesbeiana.";
 RL Arch. Biochem. Biophys. 253:241-248(1987).
 CC -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
 CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
 CC -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
 CC -!- SUBUNIT: HOMOTETRAMER.
 CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
 CC -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
 CC FAMILY.
 DR HSSP: P04179; 1MSD.
 DR InterPro: IPR001189; SOD_M1.
 DR Pfam: PF00081; sodfe; 1.
 DR PROSITE: PS00088; SOD_MN; PARTIAL.
 KW Oxidoreductase; Manganese; Mitochondrion.
 FT NON_TER 23
 SQ SEQUENCE 23 AA; 2594 MW; 5D80ED9B0E04F625 CRC64;

Query Match 17.8%; Score 24; DB 1; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 12 PDLP 15
 ||||

Db 5 PDLP 8
 RESULT 14
 CXAL_CONMA STANDARD; PRT; 14 AA.
 AC P01521;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-DEC-1992 (Rel. 24, Last annotation update)
 DE ALPHA-CONOTOXIN MI (M1).
 OS Conus magus (Magus cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=6492;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=83073458; PubMed=7149738;
 RA McIntosh J.M., Cruz L.J., Hunkapiller M.W., Gray W.R., Olivera B.M.;
 RT "Isolation and structure of a peptide toxin from the marine snail
 Conus magus.";
 RL Arch. Biochem. Biophys. 218:329-334(1982).
 RN [2]
 RP DISULFIDE BONDS.
 RX MEDLINE=84032400; PubMed=6630187;
 RA Gray W.R., Rivier J.E., Galyean R., Cruz L.J., Olivera B.M.;
 RT "Conotoxin MI. Disulfide bonding and conformational states.";
 RL J. Biol. Chem. 258:12247-12251(1983).
 RN [3]
 RP REVIEW.
 RX MEDLINE=89024586; PubMed=3052286;
 RA Gray W.R., Olivera B.M., Cruz L.J.;
 RT "Peptide toxins from venomous Conus snails.";
 RL Annu. Rev. Biochem. 57:665-700(1988).
 CC -!- FUNCTION: ALPHA-CONOTOXINS ACT ON POSTSYNAPTIC MEMBRANES, THEY
 CC BIND TO THE NICOTINIC ACETYLCHOLINE RECEPTORS (NACHR) AND THUS
 CC INHIBIT THEM.
 DR PIR: A01784; NTKN1M.
 KW Postsynaptic neurotoxin; Acetylcholine receptor inhibitor; Amidation;
 KW Venom.
 FT DISULFID 3 8
 FT DISULFID 4 14
 FT MOD_RES 14 14
 SQ SEQUENCE 14 AA; 1499 MW; DEEF1898BF5E3BD CRC64;

Query Match 17.0%; Score 23; DB 1; Length 14;
 Best Local Similarity 50.0%; Pred. No. 1.1e+03;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 CGEY 5
 ||: |
 Db 8 CGKNY 12

RESULT 15
 ID MCA2_RHOOP STANDARD; PRT; 15 AA.
 AC P56870;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE PUTATIVE MALEYLACETATE REDUCTASE II (EC 1.3.1.32) (FRAGMENT).
 OS Rhodococcus opacus (Nocardia opaca).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.
 OX NCBI_TaxID=37919;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=1CP.
 RX MEDLINE=98324954; PubMed=9657989;
 RA Seibert V., Kourbatova E.M., Golovleva L.A., Schloemann M.;
 RT "Characterization of a maleylacetate reductase encoding region from

RT Rhodococcus opacus ICP.";
RL J. Bacteriol. 180:3503-3508(1998).
CC -1- CATALYTIC ACTIVITY: 3-OXOADIPATE + NAD(P)(+) -> 2-MALEYLACETATE +
CC NAD(P)H.
CC -1- PATHWAY: 3-CHLOROCATECHOL DEGRADATION (BETA-KETOADIPATE PATHWAY).
CC THIS PATHWAY SERVES A VITAL ROLE IN THE BIODEGRADATION OF TOXIC
CC AROMATIC COMPOUNDS INTRODUCED IN THE ENVIRONMENT BOTH AS NATURAL
CC PRODUCTS AND AS INDUSTRIAL EFFLUENT.
CC -1- SIMILARITY: BELONGS TO THE IRON-CONTAINING ALCOHOL DEHYDROGENASE
CC FAMILY.
DR InterPro; IPR001670; Fe-ADH.
DR PROSITE; PS00913; ADH_IRON_1; PARTIAL.
DR PROSITE; PS00050; ADH_IRON_2; PARTIAL.
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1884 MW; 58DA90DD038F025E CRC64;

Query Match 17.0%; Score 23; DB 1; Length 15;
Best Local Similarity 37.5%; Pred. No. 1.2e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 11 HPDLPREV 18
| :||: :
Db 5 HENLPQRI 12

Search completed: March 4, 2002, 13:20:28
Job time: 615 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:21:53 ; Search time 76.45 Seconds
(without alignments)
47.833 Million cell updates/sec

Title: US-09-701-623c-84
Perfect score: 135
Sequence: 1 CQETKSTVSHPDLPREVRSIAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 7775

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mmc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 38 | 28.1 | 24 | 6 | Q9TRX6 |
| 2 | 33 | 24.4 | 21 | 2 | Q9X3C4 |
| 3 | 32 | 23.7 | 17 | 6 | Q9TR98 |
| 4 | 32 | 23.7 | 19 | 13 | O42416 |
| 5 | 31.5 | 23.3 | 17 | 4 | Q16310 |
| 6 | 30.5 | 22.6 | 21 | 4 | Q9X3J8 |
| 7 | 30 | 22.2 | 11 | 2 | Q47604 |
| 8 | 30 | 22.2 | 16 | 6 | Q9TR09 |
| 9 | 30 | 22.2 | 20 | 4 | Q9UCB8 |
| 10 | 30 | 22.2 | 20 | 6 | Q9TQ26 |
| 11 | 30 | 22.2 | 23 | 2 | P95839 |
| 12 | 29 | 21.5 | 16 | 2 | Q9R4V4 |
| 13 | 29 | 21.5 | 24 | 6 | Q9TRX4 |
| 14 | 29 | 21.5 | 25 | 8 | Q9T200 |
| 15 | 27.5 | 20.4 | 20 | 2 | O67964 |
| 16 | 27.5 | 20.4 | 20 | 2 | O67966 |
| 17 | 27.5 | 20.4 | 21 | 2 | Q9X3D8 |
| 18 | 27.5 | 20.4 | 21 | 2 | Q9X3D5 |
| 19 | 27.5 | 20.4 | 21 | 2 | Q9WVY2 |

| | | | | | | |
|----|------|------|----|----|--------|--------------------|
| 20 | 27.5 | 20.4 | 21 | 2 | Q9R2Y1 | Q9r2y1 prochloroco |
| 21 | 27.5 | 20.4 | 25 | 5 | Q9BM56 | Q9bm56 lineus sp. |
| 22 | 27 | 20.0 | 15 | 2 | O05991 | O05991 staphylococ |
| 23 | 27 | 20.0 | 18 | 2 | Q9X3E9 | Q9x3e9 prochloroco |
| 24 | 27 | 20.0 | 18 | 6 | Q28069 | Q28069 bos taurus |
| 25 | 27 | 20.0 | 19 | 2 | Q9R7I3 | Q9r7i3 prochloroco |
| 26 | 27 | 20.0 | 20 | 11 | Q9QUX8 | Q9qux8 rattus sp. |
| 27 | 27 | 20.0 | 21 | 2 | Q9WVZ3 | Q9wvz3 prochloroco |
| 28 | 27 | 20.0 | 21 | 12 | O85667 | O85667 reovirus sp |
| 29 | 27 | 20.0 | 23 | 12 | O65291 | O65291 human adeno |
| 30 | 26.5 | 19.6 | 21 | 2 | Q9X3C2 | Q9x3c2 prochloroco |
| 31 | 26 | 19.3 | 9 | 6 | Q28093 | Q28093 bos taurus |
| 32 | 26 | 19.3 | 18 | 4 | Q9UOG8 | Q9uog8 homo sapien |
| 33 | 26 | 19.3 | 18 | 11 | Q9JUM8 | Q9jnm8 rattus norv |
| 34 | 26 | 19.3 | 19 | 4 | Q16271 | Q16271 homo sapien |
| 35 | 26 | 19.3 | 20 | 2 | Q9X3M3 | Q9x3m3 prochloroco |
| 36 | 26 | 19.3 | 20 | 5 | Q9U8M9 | Q9u8m9 scaptomyza |
| 37 | 26 | 19.3 | 20 | 13 | Q9PRS2 | Q9prs2 sphenodon. |
| 38 | 26 | 19.3 | 21 | 2 | Q9X3F4 | Q9x3f4 prochloroco |
| 39 | 26 | 19.3 | 21 | 11 | Q9QVAl | Q9qva1 rattus sp. |
| 40 | 26 | 19.3 | 22 | 4 | Q13659 | Q13659 homo sapien |
| 41 | 26 | 19.3 | 22 | 8 | Q9GIA0 | Q9gia0 sargassum s |
| 42 | 26 | 19.3 | 23 | 4 | Q9H120 | Q9h120 homo sapien |
| 43 | 26 | 19.3 | 23 | 8 | Q9GIB3 | Q9gib3 sargassum c |
| 44 | 26 | 19.3 | 23 | 8 | Q9GIB2 | Q9gib2 sargassum e |
| 45 | 26 | 19.3 | 23 | 8 | Q9GIB0 | Q9gib0 sargassum b |

ALIGNMENTS

RESULT 1

Q9TRX6 PRELIMINARY; PRT; 24 AA.
AC Q9TRX6
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE GLUTAMATE DEHYDROGENASE (EC 1.4.1.3) (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=91308094; Pubmed=1854724;
RA Ozturk D.H., Colman R.F.;
RT "Identification of cysteine-319 as the target amino acid of 8-[(4-bromo-2,3-dioxobutyl)thio]adenosine 5'-triphosphate in bovine liver glutamate dehydrogenase.";
RT Biochemistry 30:7126-7134(1991).
RL SEQUENCE 24 AA; 2669 MW; D67A84AE4264E740 CRC64;

Query Match 28.1%; Score 38; DB 6; Length 24;

Best Local Similarity 50.0%; Pred. No. 80;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 3 ETYKSTVSHPDLP 14

Db 7 DTYASTIGHYDI 18

RESULT 2

Q9X3C4 PRELIMINARY; PRT; 21 AA.
ID Q9X3C4
AC Q9X3C4;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).
GN PETD.

OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanogr. 43:1615-1630(1998).
DR EMBL; AF070129; AAD20735.1; -.
FT NON_TER 21
SQ SEQUENCE 21 AA; 2318 MW; 78824B529A2C9262 CRC64;

Query Match 24.4%; Score 33; DB 2; Length 21;
Best Local Similarity 46.7%; Pred. No. 3.8e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 7 STVSHPDLPREVVR 21
Db 2 STLKPPDLRSKXKS 16
||: ||| | :|

RESULT 3
Q9TR98 PRELIMINARY; PRT; 17 AA.
AC Q9TR98:
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE ALBUMIN (FRAGMENT).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RX MEDLINE=95007849; PubMed=7923441;
RA Miller M.J., Parmelee D.C., Benjamin T., Sechi S., Dooley K.L.,
RA Kadlubar F.F.;
RT "Plasma proteins as early biomarkers of exposure to carcinogenic
RT aromatic amines";
RL Chem.-Biol. Interact. 93:221-234(1994).
DR HSP: P02768; IUOR.
SQ SEQUENCE 17 AA; 2024 MW; 1D39F70F7D23B269 CRC64;

Query Match 23.7%; Score 32; DB 6; Length 17;
Best Local Similarity 55.6%; Pred. No. 4.2e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 EYKSTVSH 11
Db 1 EYKSEIAH 9
| ||| :|

RESULT 4
O42416 PRELIMINARY; PRT; 19 AA.
AC O42416:
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-AUG-1998 (TReMBLrel. 07, Last annotation update)
DE THROMBOMUCIN (FRAGMENT).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA McNagly K.M., Petterson I., Rossi F., Flamme I., Shevchenko A.,

RA Mann M., Graf T.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y13977; CAA74310.1; -.
FT NON_TER 1
FT NON_TER 19
SQ SEQUENCE 19 AA; 2181 MW; 483C3DF97E13EC19 CRC64;

Query Match 23.7%; Score 32; DB 13; Length 19;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 HPDLP 15
Db 3 HPDLP 7
|||||

RESULT 5
Q16310 PRELIMINARY; PRT; 17 AA.
AC Q16310:
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-JAN-1999 (TReMBLrel. 09, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE GC*2 PROTEIN (FRAGMENT).
GN GC*2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95242701; PubMed=7725672;
RA Kofler A., Braun A., Jenkins T., Serjeantson S.W., Cleve H.;
RT "Characterization of mutants of the vitamin-D-binding protein/group
RT specific component: GC aborigine (1A1) from Australian aborigines and
RT South African blacks, and 2A9 from south Germany.";
RL Vox Sang. 68:50-54(1995).
DR EMBL; S77130; AAD14250.1; -.
FT NON_TER 17
SQ SEQUENCE 17 AA; 1845 MW; BB26CAD60293722C CRC64;

Query Match 23.3%; Score 31.5; DB 4; Length 17;
Best Local Similarity 40.0%; Pred. No. 5e+02;
Matches 6; Conservative 4; Mismatches 4; Indels 1; Gaps 1;

Qy 12 PD-LPREVVR 25
Db 2 PDATPKELAKLVNKC 16
|| | :|

RESULT 6
Q9X3J8 PRELIMINARY; PRT; 21 AA.
AC Q9X3J8:
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TReMBLrel. 12, Last annotation update)
DE CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).
GN PETD.
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanogr. 43:1615-1630(1998).
DR EMBL; AF070190; AAD23228.1; -.
FT NON_TER 21

SQ SEQUENCE 21 AA; 2298 MW; CC946FFCC02C854F CRC64;

Query Match 22.6%; Score 30.5; DB 2; Length 21;
Best Local Similarity 47.4%; Pred. No. 8.8e+02;

Matches 9; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 7 STVSHDPLPREVVR-IAK 24
II: III: :II:II
Db 2 STLKKPDLADPKLRSLAK 20

RESULT 7
Q47604 PRELIMINARY; PRT; 11 AA.

AC Q47604;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE C (FRAGMENT).
GN REASE.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia
OX NCBI_TaxID=562;
[1]

RP SEQUENCE FROM N.A.
RX MEDLINE=91139577; PubMed=1995588;
RA Tao T., Bourne J.C., Blumenthal R.M.;
RT "A family of regulatory genes associated with type II restriction-
modification systems.";
RL J. Bacteriol. 173:1367-1375(1991).
DR EMBL; M63621; AAA24560.1; -.
DR HSSP; P23657; 3PVI.
FT NON_TER 11
SQ SEQUENCE 11 AA; 1296 MW; 3039A71A34472AB7 CRC64;

Query Match 22.2%; Score 30; DB 2; Length 11;
Best Local Similarity 83.3%; Pred. No. 5.3e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 VSHPD 14
:IIIIII
Db 1 MSHPD 6

RESULT 8
Q9TR09 PRELIMINARY; PRT; 16 AA.

AC Q9TR09;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE PA700 SUBUNIT (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=96216387; PubMed=8621709;
RA DeMartino G.N., Proske R.J., Moomaw C.R., Strong A.A., Song X.,
RA Hisamatsu H., Tanaka K., Slaughter C.A.;
RT "Identification, purification, and characterization of a PA700-
dependent activator of the proteasome.";
RL J. Biol. Chem. 271:3112-3118(1996).
SQ SEQUENCE 16 AA; 1888 MW; FFA9A93148F3A7BF4 CRC64;

Query Match 22.2%; Score 30; DB 6; Length 16;
Best Local Similarity 60.0%; Pred. No. 7.8e+02;

Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 11 HPDLPREVVR 20
I III I I
Db 2 HIDLPNEQAR 11

RESULT 9

Q9UCE8 PRELIMINARY; PRT; 20 AA.
AC Q9UCE8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE D(TTAGGG)N-BINDING PROTEIN B39 (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=93309464; PubMed=8321232;
RA Ishikawa F., Matunis M.J., Dreyfuss G., Cech T.R.;
RT "Nuclear proteins that bind the pre-mRNA 3' splice site sequence
r(UUAG/G) and the human telomeric DNA sequence d(TTAGGG)n.";
RL Mol. Cell. Biol. 13:4301-4310(1993).
SQ SEQUENCE 20 AA; 2255 MW; C3CE1955E9A6D210 CRC64;

Query Match 22.2%; Score 30; DB 4; Length 20;
Best Local Similarity 55.6%; Pred. No. 9.9e+02;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 12 PDLPREVVR 20
I I I I I I
Db 8 PDTPEXIR 16

RESULT 10

Q9TQZ6 PRELIMINARY; PRT; 20 AA.
AC Q9TQZ6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE ALBUMIN (FRAGMENT).
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE.
RX MEDLINE=96273610; PubMed=8690030;
RA Nicolas M.G., Fujiki K., Murayama K., Suzuki M.T., Mineki R.,
RA Hayakawa M., Yoshikawa Y., Cho F., Kanai A.;
RT "Studies on the mechanism of early onset macular degeneration in
Cynomolgus (Macaca fascicularis) monkeys. I. Abnormal concentrations
of two proteins in the retina.";
RL Exp. Eye Res. 62:211-219(1996).
SQ SEQUENCE 20 AA; 2411 MW; 5F1A6AEB5918F777 CRC64;

Query Match 22.2%; Score 30; DB 6; Length 20;
Best Local Similarity 55.6%; Pred. No. 9.9e+02;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 ETKYSTVSH 11
:I:II I:I
Db 11 DTHKSEVAH 19

RESULT 11

```
P95839
ID P95839          PRELIMINARY:      PRT;      23 AA.
AC P95839
DT 01-MAY-1997 (TReMBLrel. 03, Created)
DT 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE ORF56 (FRAGMENT).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COL;
RX MEDLINE=96427339; PubMed=8830703;
RA Wu S., de Lencastre H., Tomasz A.;
RT "Sigma-B, a putative operon encoding alternate sigma factor of
RT Staphylococcus aureus RNA polymerase: molecular cloning and DNA
RT sequencing.";
RL J. Bacteriol. 178:6036-6042(1996).
DR EMBL: Y09929; CAA71063.1;
FT NON_TER 1
SQ SEQUENCE 23 AA; 2541 MW; 7F47717B1767D34F CRC64;

Query Match 22.2%; Score 30; DB 2; Length 23;
Best Local Similarity 60.0%; Pred. No. 1.1e+03;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGETYKSTVS 10
   | | | | |
Db 11 CNETYLSNS 20

RESULT 12
Q9R4V4
ID Q9R4V4          PRELIMINARY:      PRT;      16 AA.
AC Q9R4V4
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE AMADORI PRODUCT BINDING PROTEIN (FRAGMENT).
OS Pseudomonas (fluorescent pseudomonads).
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae.
OX NCBI_TaxID=286;
RN [1]
RP SEQUENCE.
RX MEDLINE=95050465; PubMed=7961640;
RA Gerhardsinger C., Taneda S., Marion M.S., Monnier V.M.;
RT "Isolation, purification, and characterization of an Amadori product
RT binding protein from a Pseudomonas sp. soil strain.";
RL J. Biol. Chem. 269:27297-27302(1994).
SQ SEQUENCE 16 AA; 1619 MW; 7330EA1CD1193CD6 CRC64;

Query Match 21.5%; Score 29; DB 2; Length 16;
Best Local Similarity 45.5%; Pred. No. 1.1e+03;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 7 STVSHPDLPRE 17
   : | : | |
Db 3 AVVAEPDAPAE 13

RESULT 13
Q9TRX4
ID Q9TRX4          PRELIMINARY:      PRT;      24 AA.
AC Q9TRX4
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE GLUTAMATE DEHYDROGENASE (EC 1.4.1.3) (FRAGMENT).
OS Bos taurus (bovine).

us-09-701-623c-84.closed.rspt

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Ox NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=91308094; PubMed=1854724;
RA Ozturk D.H., Colman R.F.;
RT "Identification of cysteine-319 as the target amino acid of 8-[(4-
RT bromo-2,3-dioxobutyl)thio]adenosine 5'-triphosphate in bovine liver
RT glutamate dehydrogenase.";
RL Biochemistry 30:7126-7134(1991).
SQ SEQUENCE 24 AA; 2626 MW; DE4ED4B54264E740 CRC64;

Query Match 21.5%; Score 29; DB 6; Length 24;
Best Local Similarity 41.7%; Pred. No. 1.7e+03;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 3 ETYKSTVSHPDLP 14
   : | | | | :
Db 7 DTYASTICXYDI 18

RESULT 14
Q9T2Q0
ID Q9T2Q0          PRELIMINARY:      PRT;      25 AA.
AC Q9T2Q0
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE MANGANESE SUPEROXIDE DISMUTASE (FRAGMENT).
OS Solanum tuberosum (Potato).
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE.
RX MEDLINE=95392474; PubMed=7545053;
RA Fester T., Schuster W.;
RT "Potato mitochondrial manganese superoxide dismutase is an RNA-binding
RT protein.";
RL Biochem. Mol. Biol. Int. 36:67-75(1995).
CC -1- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) -> O(2) + H(2)O(2).
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC FAMILY.
DR HSPSP; P04179; IAP6.
DR InterPro; IPR001189; SOD_MI.
DR Pfam; PF00081; sodfe; 1.
DR ProDom; PD000475; SOD_MI; 1.
KW Manganese; Oxidoreductase.
SQ SEQUENCE 25 AA; 2713 MW; 45E0DA2EB3ADB3C3 CRC64;

Query Match 21.5%; Score 29; DB 8; Length 25;
Best Local Similarity 75.0%; Pred. No. 1.8e+03;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 8 TVSHPDLP 15
   | | | | |
Db 3 TFSLPDLP 10

RESULT 15
O67964
ID O67964          PRELIMINARY:      PRT;      20 AA.
AC O67964
DT 01-AUG-1998 (TReMBLrel. 07, Created)
```

DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
 DE 01-NOV-1998 (Tremblrel. 08, Last annotation update)
 DE BG/F COMPLEX SUBUNIT IV (FRAGMENT).
 GN PETD.
 OS Prochlorococcus sp.
 OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
 OC Prochlorococcus.
 OX NCBI_TaxID=1220;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CCMP1378;
 RX MEDLINE=98123172; PubMed=9452521;
 RA Urbach E., Scanlan D.J., Distel D.L., Waterbury J.B., Chisholm S.W.;
 RT "Rapid diversification of marine picophytoplankton with dissimilar
 -RT light-harvesting structures inferred from sequences of Prochlorococcus
 RT and Synechococcus (Cyanobacteria).";
 RL J. Mol. Evol. 46:188-201(1998).
 DR EMBL: AF001488; AAC05623.1; -;
 FT NON-TER 20 20
 SQ SEQUENCE 20 AA; 2241 MW; C46FEA1A34E54FDC CRC64;

Query Match 20.4%; Score 27.5; DB 2; Length 20;
 Best Local Similarity 42.1%; Pred. No. 2.3e+03;
 Matches 8; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

OY 7 STVSHPDLPREVRS-IAK 24
 Db 2 STLKKPDLSDPKLRKLA 20
 ||: ||| :||:
 ||: ||| :||:

Search completed: March 4, 2002, 13:21:53
 Job time: 654 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:10:53 ; Search time 37.41 Seconds
(without alignments)
15.038 Million cell updates/sec

Title: US-09-701-623c-84
Perfect score: 135
Sequence: 1 CGETYKSTVSHPDLPREVVRSIAC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 123821

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 94 | 69.6 | 25 | 3 | US-09-100-414B-95 |
| 2 | 94 | 69.6 | 25 | 4 | US-09-303-323-95 |
| 3 | 58 | 43.0 | 22 | 2 | US-08-232-539D-19 |
| 4 | 58 | 43.0 | 24 | 2 | US-08-232-539D-20 |
| 5 | 35 | 25.9 | 18 | 2 | US-09-017-205-5 |
| 6 | 35 | 25.9 | 21 | 4 | US-09-077-991-5 |
| 7 | 32 | 23.7 | 18 | 4 | US-09-177-249-298 |
| 8 | 31 | 23.0 | 11 | 1 | US-08-269-441A-15 |
| 9 | 31 | 23.0 | 13 | 4 | US-09-078-173A-23 |
| 10 | 31 | 23.0 | 14 | 4 | US-09-078-173A-24 |
| 11 | 31 | 23.0 | 24 | 1 | US-07-976-358-24 |
| 12 | 31 | 23.0 | 25 | 1 | US-07-976-358-15 |
| 13 | 31 | 23.0 | 25 | 1 | US-07-976-358-18 |
| 14 | 31 | 23.0 | 25 | 1 | US-07-976-358-21 |
| 15 | 30 | 22.2 | 16 | 4 | US-09-273-565-53 |
| 16 | 30 | 22.2 | 22 | 1 | US-08-460-874A-37 |
| 17 | 30 | 22.2 | 22 | 2 | US-08-388-883B-37 |
| 18 | 30 | 22.2 | 22 | 4 | US-08-462-211A-37 |
| 19 | 30 | 22.2 | 25 | 1 | US-07-976-358-17 |
| 20 | 29 | 21.5 | 10 | 1 | US-08-041-774-1 |
| 21 | 29 | 21.5 | 10 | 4 | US-08-530-340-7 |
| 22 | 29 | 21.5 | 14 | 3 | US-08-405-647B-9 |
| 23 | 29 | 21.5 | 14 | 4 | US-08-985-499-9 |
| 24 | 29 | 21.5 | 14 | 5 | PCT-US96-03180-9 |
| 25 | 29 | 21.5 | 14 | 6 | 5164482-18 |
| 26 | 29 | 21.5 | 15 | 2 | US-09-049-577-4 |
| 27 | 29 | 21.5 | 15 | 4 | US-09-390-598-4 |

| | | | | | | |
|----|------|------|----|---|--------------------|-------------------|
| 28 | 29 | 21.5 | 18 | 2 | US-09-017-205-6 | Sequence 6, Appl |
| 29 | 29 | 21.5 | 20 | 2 | US-08-564-972-34 | Sequence 34, Appl |
| 30 | 29 | 21.5 | 21 | 1 | US-08-786-748A-53 | Sequence 53, Appl |
| 31 | 29 | 21.5 | 21 | 1 | US-08-786-748A-67 | Sequence 67, Appl |
| 32 | 29 | 21.5 | 21 | 2 | US-08-932-682-53 | Sequence 53, Appl |
| 33 | 29 | 21.5 | 21 | 2 | US-08-932-682-67 | Sequence 67, Appl |
| 34 | 29 | 21.5 | 22 | 4 | US-09-162-934-12 | Sequence 12, Appl |
| 35 | 29 | 21.5 | 23 | 4 | US-09-253-396A-218 | Sequence 218, App |
| 36 | 29 | 21.5 | 24 | 1 | US-08-786-748A-60 | Sequence 60, Appl |
| 37 | 29 | 21.5 | 24 | 2 | US-08-932-682-60 | Sequence 60, Appl |
| 38 | 29 | 21.5 | 25 | 4 | US-09-248-588-56 | Sequence 56, Appl |
| 39 | 29 | 21.5 | 25 | 4 | US-09-488-799-45 | Sequence 45, Appl |
| 40 | 28.5 | 21.1 | 18 | 2 | US-08-797-842-8 | Sequence 8, Appl |
| 41 | 28 | 20.7 | 8 | 2 | US-08-686-599A-11 | Sequence 11, Appl |
| 42 | 28 | 20.7 | 10 | 3 | US-09-139-762A-68 | Sequence 68, Appl |
| 43 | 28 | 20.7 | 13 | 1 | US-08-435-019-10 | Sequence 10, Appl |
| 44 | 28 | 20.7 | 14 | 2 | US-08-332-562A-59 | Sequence 59, Appl |
| 45 | 28 | 20.7 | 15 | 1 | US-08-221-583-45 | Sequence 45, Appl |

ALIGNMENTS

RESULT 1
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-95

Query Match 69.6% Score 94; DB 3; Length 25;
Best Local Similarity 64.0%; Pred. No. 7.4e-09;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHPDLPREVVRSIAC 25
|||||:|:|:|:|:|:|:|:|:|

DB 1 CGETYQSNVTHPLPALMRSTTKC 25
|||||:|:|:|:|:|:|:|:|:

RESULT 2
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang YI
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 69.6%; Score 94; DB 4; Length 25;
Best Local Similarity 64.0%; Pred. No. 7.4e-09;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHPDLPREVRSIARC 25
Db 1 CGETYQSRVTHPLPRALMRSTTKC 25

RESULT 3
US-08-232-539D-19
; Sequence 19, Application US/08232539D
; Patent No. 5965709
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; APPLICANT: Jardieu, Paula M.
; TITLE OF INVENTION: Ige Antagonists
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-232-539D-19

Query Match 43.0%; Score 58; DB 2; Length 22;
Best Local Similarity 55.6%; Pred. No. 0.0046;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 2 GETYKSTVSHPDLPREV 19
Db 5 GETYQSRVTHPLPRALM 22

RESULT 4
US-08-232-539D-20
; Sequence 20, Application US/08232539D
; Patent No. 5965709
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; APPLICANT: Jardieu, Paula M.
; TITLE OF INVENTION: Ige Antagonists
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881

; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-232-539D-20

Query Match 43.0%; Score 58; DB 2; Length 24;
Best Local Similarity 55.6%; Pred. No. 0.0051;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 2 GETYKSTVSHPDLPREV 19
| | | | | :
Db 7 GETYQCRVTHPHLPALM 24

RESULT 5
US-09-017-205-5
; Sequence 5, Application US/09017205
; Patent No. 5965357
; GENERAL INFORMATION:
; APPLICANT: Marsden, Howard S
; TITLE OF INVENTION: PEPTIDE STRUCTURES AND THEIR USE IN
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
; STREET: 8th Floor, 1100 No. 5965357th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,205
; FILING DATE: 02-FEB-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mitchard, Leonard C
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 604-436
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4000
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide from HSV-2 glycoprotein G
; FRAGMENT TYPE: internal
US-09-017-205-5

Query Match 25.9%; Score 35; DB 2; Length 18;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 13 DLPREVVR 20
| | | | |
Db 10 DLPREVVR 17

RESULT 6
US-09-077-991-5
; Sequence 5, Application US/09077991
; Patent No. 6207375
; GENERAL INFORMATION:

; APPLICANT: Subramaniam, M.
; APPLICANT: Spelsberg, T.C.
; APPLICANT: Roche, P.C.
; TITLE OF INVENTION: TGF-Beta inducible early factor-1
; TITLE OF INVENTION: (TIEF-1) and a method to detect breast cancer
; FILE REFERENCE: 150.1570S2
; CURRENT APPLICATION NUMBER: US/09/077,991
; CURRENT FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: PCT/US96/19555
; EARLIER FILING DATE: 1996-12-11
; EARLIER APPLICATION NUMBER: US 08/570,227
; EARLIER FILING DATE: 1995-12-11
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-077-991-5

Query Match 25.9%; Score 35; DB 4; Length 21;
Best Local Similarity 57.1%; Pred. No. 24;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 3 ETYKSTVSHPDLP 16
| | | | |
Db 3 EEKSPVSAKPK 16

RESULT 7
US-09-177-249-298
; Sequence 298, Application US/09177249
; Patent No. 6229064
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Ohad, Nir
; APPLICANT: Kiyosue, Tomohiro
; APPLICANT: Yadegari, Ramin
; APPLICANT: Margossian, Linda
; APPLICANT: Harada, John
; APPLICANT: Goldberg, Robert B.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit
; TITLE OF INVENTION: Development in Plants
; FILE REFERENCE: 023070-086120US
; CURRENT APPLICATION NUMBER: US/09/177,249
; CURRENT FILING DATE: 1998-10-22
; EARLIER APPLICATION NUMBER: US 09/071,838
; EARLIER FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 298
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-09-177-249-298

Query Match 23.7%; Score 32; DB 4; Length 18;
Best Local Similarity 50.0%; Pred. No. 61;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 16 REVVRSIAC 25
: | | | : | |
Db 2 KSVVRNVQKC 11

RESULT 8
US-08-269-441A-15
; Sequence 15, Application US/08269441A
; Patent No. 5552529
; GENERAL INFORMATION:

APPLICANT: Rearden, Ann
TITLE OF INVENTION: A NOVEL AUTOANTIGEN, PINCH
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/269,441A
FILING DATE: 30-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haile Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07257/009001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..11
US-08-269-441A-15

Query Match 23.0% Score 31; DB 1; Length 11;
Best Local Similarity 45.5%; Pred. No. 49;
Matches 5; Conservative 2; Mismatches 4; Indels 4; Gaps 0;

Qy 5 YKSTVSHPDLP 15
Db 1 FKPPYHPDVP 11

RESULT 9
US-09-078-173A-23
Sequence 23, Application US/09078173A
Patent No. 6200794
GENERAL INFORMATION:
APPLICANT: Ian M. Whitehead
APPLICANT: Alan Slusarenko
APPLICANT: Duncan Gaskins
APPLICANT: Alan Brash
APPLICANT: Nathalie Tijet
TITLE OF INVENTION: GUAVA (PSIDIUM GUAJAVA) 13-HYDROPEROXIDE
FILE REFERENCE: 06027.0001
CURRENT APPLICATION NUMBER: US/09/078,173A
CURRENT FILING DATE: 1998-05-13
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 33
TYPE: PRT
ORGANISM: Psidium Guajava (guava)
US-09-078-173A-23

Query Match 23.0% Score 31; DB 4; Length 13;

Best Local Similarity 46.2%; Pred. No. 60;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
Qy 4 TYKSTVSHPDLP 16
Db 1 TYPPLSPSPSPR 13

RESULT 10
US-09-078-173A-24
Sequence 24, Application US/09078173A
Patent No. 6200794
GENERAL INFORMATION:
APPLICANT: Ian M. Whitehead
APPLICANT: Alan Slusarenko
APPLICANT: Duncan Gaskins
APPLICANT: Alan Brash
APPLICANT: Nathalie Tijet
TITLE OF INVENTION: GUAVA (PSIDIUM GUAJAVA) 13-HYDROPEROXIDE
FILE REFERENCE: 06027.0001
CURRENT APPLICATION NUMBER: US/09/078,173A
CURRENT FILING DATE: 1998-05-13
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 24
LENGTH: 14
TYPE: PRT
ORGANISM: Psidium Guajava (guava)
US-09-078-173A-24

Query Match 23.0% Score 31; DB 4; Length 14;
Best Local Similarity 46.2%; Pred. No. 65;
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 4 TYKSTVSHPDLP 16
Db 1 TYPPLSPSPSPR 13

RESULT 11
US-07-976-358-24
Sequence 24, Application US/07976358
Patent No. 5445932
GENERAL INFORMATION:
APPLICANT: FIELDS, Howard A.
APPLICANT: KHUDYAKOV, Yuri
APPLICANT: FAVOROV, Michael
TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
TITLE OF INVENTION: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/976,358
FILING DATE: 19921117
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/134 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-976-358-24

Query Match 23.0%; Score 31; DB 1; Length 24;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 5 YKSTVSHDPLPREV 18
| | | | |
Db 7 YSQFTPLSLPRV 20

RESULT 12

US-07-976-358-15
Sequence 15, Application US/07976358
Patent No. 5445932

GENERAL INFORMATION:
APPLICANT: FIELDS, Howard A.
APPLICANT: KHUYAKOV, Yuri
TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
NUMBER OF SEQUENCES: 24
CURRENT APPLICATION DATA: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/976,358
FILING DATE: 19921117

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/134 NIHD
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-976-358-15

Qy 5 YKSTVSHDPLPREV 18
| | | | |
Db 8 YSQFTPLSLPRV 21

US-07-976-358-15
Sequence 15, Application US/07976358
Patent No. 5445932

GENERAL INFORMATION:
APPLICANT: FIELDS, Howard A.
APPLICANT: KHUYAKOV, Yuri
TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
NUMBER OF SEQUENCES: 24
CURRENT APPLICATION DATA: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

Query Match 23.0%; Score 31; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 5 YKSTVSHDPLPREV 18
| | | | |
Db 8 YSQFTPLSLPRV 21

RESULT 13

US-07-976-358-18
Sequence 18, Application US/07976358

Patent No. 5445932
GENERAL INFORMATION:
APPLICANT: FIELDS, Howard A.
APPLICANT: KHUYAKOV, Yuri
TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
NUMBER OF SEQUENCES: 24
CURRENT APPLICATION DATA: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

Query Match 23.0%; Score 31; DB 1; Length 25;
Best Local Similarity 52.9%; Pred. No. 1.3e+02;
Matches 9; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 5 YKSTVSHDPLPREV 21
| | | | |
Db 8 YSSHPLPLPRV 24

RESULT 14

US-07-976-358-21
Sequence 21, Application US/07976358
Patent No. 5445932

GENERAL INFORMATION:
APPLICANT: FIELDS, Howard A.
APPLICANT: KHUYAKOV, Yuri
TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
NUMBER OF SEQUENCES: 24
CURRENT APPLICATION DATA: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109

Query Match 23.0%; Score 31; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 5 YKSTVSHDPLPREV 18
| | | | |
Db 8 YSQFTPLSLPRV 21

RESULT 13

US-07-976-358-18
Sequence 18, Application US/07976358

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/134 NIHD
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US-07-976-358-21

Query Match 23.0%; Score 31; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 YKSTVSHPDLPREV 18
| | : | | | |
Db 8 YSQSTLPSLPVV 21

RESULT 15
US-09-273-565-53
; Sequence 53, Application US/09273565A
; Patent No. 6166.90
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, TSUTOMU
; APPLICANT: WATANABE, TAKESHI
; APPLICANT: HORIE, MASATO
; TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
; FILE REFERENCE: Q-53599
; CURRENT APPLICATION NUMBER: US/09/273,565A
; CURRENT FILING DATE: 1999-03-22
; EARLIER APPLICATION NUMBER: 09/055,699
; EARLIER FILING DATE: 1998-04-07
; EARLIER APPLICATION NUMBER: 08/820,170
; EARLIER FILING DATE: 1997-03-19
; EARLIER APPLICATION NUMBER: JP 63410/1996
; EARLIER FILING DATE: 1996-03-19
; EARLIER APPLICATION NUMBER: JP 69163/1997
; EARLIER FILING DATE: 1997-03-05
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Bovine sp.
US-09-273-565-53

Query Match 22.2%; Score 30; DB 4; Length 16;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 11 HPDLPREVVR 20
| | | | |
Db 2 HIDLPEQAR 11

Search completed: March 4, 2002, 13:10:53
Job time: 300 sec

..y

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:05:53 ; Search time 37.41 Seconds
(without alignments)
15.038 Million cell updates/sec

Title: US-09-701-623c-5

Perfect score: 140
Sequence: 1 CGETYQSRVTHPLPALMRSTTKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 123821

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 140 | 100.0 | 25 | 3 | US-09-100-414B-95 |
| 2 | 140 | 100.0 | 25 | 4 | US-09-303-323-95 |
| 3 | 93 | 66.4 | 22 | 2 | US-08-232-539D-19 |
| 4 | 93 | 66.4 | 24 | 2 | US-08-232-539D-20 |
| 5 | 35 | 25.0 | 24 | 1 | US-07-976-358-24 |
| 6 | 35 | 25.0 | 25 | 1 | US-07-976-358-15 |
| 7 | 35 | 25.0 | 25 | 1 | US-07-976-358-18 |
| 8 | 35 | 25.0 | 25 | 1 | US-07-976-358-21 |
| 9 | 34 | 24.3 | 25 | 1 | US-07-976-358-17 |
| 10 | 33 | 23.6 | 18 | 2 | US-09-017-205-9 |
| 11 | 32.5 | 23.2 | 15 | 2 | US-08-671-094B-8 |
| 12 | 32.5 | 23.2 | 16 | 1 | US-08-077-797A-14 |
| 13 | 32.5 | 23.2 | 16 | 5 | PCT-US94-01238-14 |
| 14 | 32.5 | 23.2 | 17 | 1 | US-08-311-307B-9 |
| 15 | 32 | 22.9 | 24 | 1 | US-07-976-358-23 |
| 16 | 32 | 22.9 | 25 | 1 | US-07-976-358-16 |
| 17 | 31 | 22.1 | 17 | 1 | US-08-218-025A-186 |
| 18 | 31 | 22.1 | 18 | 2 | US-09-017-205-52 |
| 19 | 31 | 22.1 | 21 | 4 | US-09-077-991-5 |
| 20 | 30 | 21.4 | 13 | 1 | US-08-466-285-6 |
| 21 | 30 | 21.4 | 16 | 4 | US-08-602-999A-226 |
| 22 | 30 | 21.4 | 18 | 2 | US-08-480-190-79 |
| 23 | 30 | 21.4 | 18 | 2 | US-08-488-379-79 |
| 24 | 30 | 21.4 | 18 | 5 | PCT-US93-07545-79 |
| 25 | 30 | 21.4 | 24 | 3 | US-08-592-500-27 |
| 26 | 30 | 21.4 | 24 | 3 | US-08-592-500-35 |
| 27 | 30 | 21.4 | 24 | 3 | US-08-195-006-27 |

| | | | | | | |
|----|------|------|----|---|--------------------|-------------------|
| 28 | 30 | 21.4 | 24 | 3 | US-08-195-006-35 | Sequence 35, Appl |
| 29 | 30 | 21.4 | 24 | 5 | PCT-US94-07644A-27 | Sequence 27, Appl |
| 30 | 30 | 21.4 | 24 | 5 | PCT-US94-07644A-35 | Sequence 35, Appl |
| 31 | 29.5 | 21.1 | 14 | 2 | US-08-726-306A-82 | Sequence 82, Appl |
| 32 | 29.5 | 21.1 | 18 | 4 | US-09-146-580-3 | Sequence 3, Appl |
| 33 | 29.5 | 21.1 | 24 | 2 | US-08-637-759B-313 | Sequence 313, App |
| 34 | 29.5 | 21.1 | 24 | 3 | US-08-871-355A-313 | Sequence 313, App |
| 35 | 29 | 20.7 | 11 | 1 | US-08-548-540-160 | Sequence 160, App |
| 36 | 29 | 20.7 | 11 | 5 | PCT-US96-09809-160 | Sequence 160, App |
| 37 | 29 | 20.7 | 18 | 4 | US-08-602-999A-371 | Sequence 371, App |
| 38 | 29 | 20.7 | 18 | 5 | PCT-US92-07813-7 | Sequence 7, Appl |
| 39 | 29 | 20.7 | 19 | 1 | US-08-116-733-12 | Sequence 12, Appl |
| 40 | 29 | 20.7 | 19 | 1 | US-08-469-615-9 | Sequence 9, Appl |
| 41 | 29 | 20.7 | 19 | 1 | US-08-466-763-9 | Sequence 9, Appl |
| 42 | 29 | 20.7 | 19 | 2 | US-08-411-142A-9 | Sequence 9, Appl |
| 43 | 29 | 20.7 | 20 | 1 | US-08-440-861-37 | Sequence 37, Appl |
| 44 | 29 | 20.7 | 20 | 2 | US-08-564-972-34 | Sequence 34, Appl |
| 45 | 29 | 20.7 | 22 | 2 | US-08-310-912A-36 | Sequence 36, Appl |

ALIGNMENTS

RESULT 1
US-09-100-414B-95
Sequence 35, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 95:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-95

Query Match 100.0%; Score 140; DB 3; Length 25;
Best Local Similarity 100.0%; Pred. No. 1e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGETYQSRVTHPLPALMRSTTKC 25
DB 1 CGETYQSRVTHPLPALMRSTTKC 25

```

; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
;
US-08-232-539D-19

```

```

Query Match 66.4%; Score 93; DB 2; Length 22;
Best Local Similarity 94.4%; Pred. No. 5.1e-09;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      2 GETYQSRVTHPLPRALM 19
        ||||| ||||| |||||
Db       5 GETYQCRVTHPLPRALM 22

```

```

RESULT      4
US-08-232-539D-20
; Sequence 20, Application US/08232539D
; Patent No. 5965709
; GENERAL INFORMATION:
; APPLICANT: Presta, Leonard G.
; APPLICANT: Jardieu, Paula M.
; TITLE OF INVENTION: Ige Antagonists
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,539D
; FILING DATE: 21-Apr-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/178583
; FILING DATE: 07-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/744768
; FILING DATE: 14-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P0718P3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881

```

; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-232-539D-20

Query Match 66.4%; Score 93; DB 2; Length 24;
Best Local Similarity 94.4%; Pred. No. 5.7e-09;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPLPRALM 19
| | | | | | | | | | | | | | | | | | | | | |
Db 7 GETYQCRVTHPLPRALM 24

RESULT 5

US-07-976-358-24
; Sequence 24, Application US/07976358
; Patent No. 5445932
; GENERAL INFORMATION:
; APPLICANT: FIELDS, Howard A.
; APPLICANT: KHUDYAKOV, Yuri
; APPLICANT: FAVOROV, Michael
; TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
; TITLE OF INVENTION: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/976,358
; FILING DATE: 19921117
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/134 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-976-358-24

Query Match 25.0%; Score 35; DB 1; Length 24;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 YQSRVTHPLPRAL 18
| | | | | | | | | | | | | | | | | | | | | |
Db 7 YSSQPTLPSLPRV 20

RESULT 6

US-07-976-358-15
; Sequence 15, Application US/07976358
; Patent No. 5445932
; GENERAL INFORMATION:

; APPLICANT: FIELDS, Howard A.
; APPLICANT: KHUDYAKOV, Yuri
; APPLICANT: FAVOROV, Michael
; TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
; TITLE OF INVENTION: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/976,358
; FILING DATE: 19921117
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/134 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-976-358-15

Query Match 25.0%; Score 35; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 YQSRVTHPLPRAL 18
| | | | | | | | | | | | | | | | | | | | | |
Db 8 YSSQPTLPSLPRV 21

RESULT 7

US-07-976-358-18
; Sequence 18, Application US/07976358
; Patent No. 5445932
; GENERAL INFORMATION:
; APPLICANT: FIELDS, Howard A.
; APPLICANT: KHUDYAKOV, Yuri
; APPLICANT: FAVOROV, Michael
; TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
; TITLE OF INVENTION: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/976,358
; FILING DATE: 19921117
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

Query Match 25.0%; Score 35; DB 1; Length 24;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 YQSRVTHPLPRAL 18
| | | | | | | | | | | | | | | | | | | | | |
Db 7 YSSQPTLPSLPRV 20

RESULT 6

US-07-976-358-15
; Sequence 15, Application US/07976358
; Patent No. 5445932
; GENERAL INFORMATION:

NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/134 NIHD
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-976-358-18

Query Match 25.0%; Score 35; DB 1; Length 25;
Best Local Similarity 47.1%; Pred. No. 22;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 5 YOSRVTHPLPALMRS 21
Db 8 YSSQTLPLPRV 24

RESULT 8
US-07-976-358-21
Sequence 21, Application US/07976358
Patent No. 5445932
GENERAL INFORMATION:
APPLICANT: FIELDS, Howard A.
APPLICANT: KHUYAKOV, Yuri
TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
TITLE OF INVENTION: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/976,358
FILING DATE: 19921117
CLASSIFICATION: 435
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/134 NIHD
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-976-358-21

Query Match 25.0%; Score 35; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 22;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 5 YOSRVTHPLPAL 18
Db 8 YSSQTLPLPRV 21

RESULT 9
US-07-976-358-17
Sequence 17, Application US/07976358
Patent No. 5445932
GENERAL INFORMATION:
APPLICANT: FIELDS, Howard A.
APPLICANT: KHUYAKOV, Yuri
TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
TITLE OF INVENTION: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/976,358
FILING DATE: 19921117
CLASSIFICATION: 435
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/134 NIHD
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
US-07-976-358-17

Query Match 24.3%; Score 34; DB 1; Length 25;
Best Local Similarity 50.0%; Pred. No. 33;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 5 YOSRVTHPLPAL 18
Db 8 YSSQTLPLPRV 21

RESULT 10
US-09-017-205-9
Sequence 9, Application US/09017205
Patent No. 5965357
GENERAL INFORMATION:
APPLICANT: Marsden, Howard S.
TITLE OF INVENTION: PEPTIDE STRUCTURES AND THEIR USE IN
TITLE OF INVENTION: DIAGNOSIS OF HERPES SIMPLEX VIRUS TYPE 2
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon & Vanderhye PC
STREET: 8th Floor, 1100 No. 5965357th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: USA
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,205
; FILING DATE: 02-FEB-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mitchard, Leonard C
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 604-436
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)816-4000
; TELEFAX: (703)816-4100
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide from HSV-2 glycoprotein G
; FRAGMENT TYPE: internal
; US-09-017-205-9

Query Match 23.6%; Score 33; DB 2; Length 18;
Best Local Similarity 53.8%; Pred. No. 32;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 TYQSRVTHPLPR 16
   11 :|||: |
Db 6 TYAARVTVRLTR 18

RESULT 11
US-08-671-094B-8
; Sequence 8, Application US/08671094B
; Patent No. 5912232
; GENERAL INFORMATION:
; APPLICANT: Talmadge, James E.
; TITLE OF INVENTION: Anti-inflammatory Polypeptide
; TITLE OF INVENTION: Antagonists of Human Interleukin-8
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/671,094B
; FILING DATE: 28-JUN-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagan, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 53086FWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100
; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
```

```
; NAME/KEY: Modified-site
; LOCATION: 4..6
; OTHER INFORMATION: /note= "The cysteine residues
; OTHER INFORMATION: may be substituted with aminobutyric acid, homocysteine or
; OTHER INFORMATION: diaminosuberic acid."
; US-08-671-094B-8

Query Match 23.2%; Score 32.5; DB 2; Length 15;
Best Local Similarity 53.8%; Pred. No. 31;
Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 CGETYQSRVTHPH 13
   1 :|||: |||
Db 4 CIRTY-SKPFHPH 15

RESULT 12
US-08-077-797A-14
; Sequence 14, Application US/08077797A
; Patent No. 5679548
; GENERAL INFORMATION:
; APPLICANT: Barbas, Carlos F.
; APPLICANT: Rosenblum, Jonathan
; APPLICANT: Lerner, Richard A.
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE METAL
; TITLE OF INVENTION: BINDING SITES AND COMPOSITIONS THEREOF
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5679548th Torrey Pines Road, TPC8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/077,797A
; FILING DATE: 14-JUN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/012,566
; FILING DATE: 02-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: SCRI276P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-077-797A-14

Query Match 23.2%; Score 32.5; DB 1; Length 16;
Best Local Similarity 42.9%; Pred. No. 34;
Matches 6; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

QY 2 GETYQSRVTHPLP 15
   1:|:::| |||
Db 1 GDTHRGHLRH-HLP 13
```

RESULT 13
PCT-US94-01238-14
; Sequence 14, Application PC/TUS9401238
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE METAL
; TITLE OF INVENTION: BINDING SITES AND COMPOSITIONS THEREOF
; NUMBER OF SEQUENCES: 65
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01238
; FILING DATE: 01-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/077,797
; FILING DATE: 14-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/012,566
; FILING DATE: 02-FEB-1993
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
PCT-US94-01238-14

Query Match 23.2%; Score 32.5; DB 5; Length 16;
Best Local Similarity 42.9%; Pred. No. 34;
Matches 6; Conservative 4; Mismatches 3; Indels 1; Gaps 1;

Qy 2 GETYQSEVTHPLP 15
I : I : I : I : I : I :
Db 1 GDTHRHLP-HLP 13

RESULT 14
US-08-311-307B-9
; Sequence 9, Application US/08311307B
; Patent No. 5627156
; GENERAL INFORMATION:
; APPLICANT: Talmadge, James E.
; TITLE OF INVENTION: Polypeptide Agonist Derived From Human
; TITLE OF INVENTION: Interleukin-8
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dann, Dorfman, Herrell and Skillman
; STREET: 1601 Market Street
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,307B
; FILING DATE: 23-SEP-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hagan, Patrick J.
; REGISTRATION NUMBER: 27,643
; REFERENCE/DOCKET NUMBER: 63085
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215)563-4100

; TELEFAX: (215)563-4044
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 4..6
; OTHER INFORMATION: /note= "The cysteine residues may
; OTHER INFORMATION: be substituted with aminobutyric acid, homocysteine or
; OTHER INFORMATION: diaminosuberic acid."
US-08-311-307B-9

Query Match 23.2%; Score 32.5; DB 1; Length 17;
Best Local Similarity 53.8%; Pred. No. 36;
Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Qy 1 CGETYQSRVTHPH 13
I : I : I : I : I : I :
Db 6 CIKTY-SKPFPH 17

RESULT 15
US-07-976-358-23
; Sequence 23, Application US/07976358
; Patent No. 5445932
; GENERAL INFORMATION:
; APPLICANT: FIELDS, Howard A.
; APPLICANT: KHUDYAKOV, Yuri
; TITLE OF INVENTION: FAVOROV, Michael
; TITLE OF INVENTION: METHOD FOR DETECTION OF A NEW MARKER
; TITLE OF INVENTION: ASSOCIATED WITH HEPATITIS DELTA VIRUS INFECTION
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/976,358
; FILING DATE: 19921117
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/134 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
US-07-976-358-23

Query Match 22.9%; Score 32; DB 1; Length 24;
Best Local Similarity 42.9%; Pred. No. 67;
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 5 YQSRVTHPLRAL 18
| | : | | | :
Db 7 YSSQIRPFLPRV 20

Search completed: March 4, 2002, 13:10:52
Job time: 299 sec

THIS PAGE BLANK (unfiled)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:52:18 ; Search time 38.04 Seconds
(without alignments)
14.789 Million cell updates/sec

Title: US-09-701-623C-5

Perfect score: 140

Sequence: 1 CGETYQSRVTHPLPRALMRSTTKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1 | 140 | 100.0 | 25 | 3 US-09-100-414B-95 | Sequence 95, Appl |
| 2 | 140 | 100.0 | 25 | 4 US-09-303-323-95 | Sequence 95, Appl |
| 3 | 140 | 100.0 | 42 | 3 US-09-100-414B-98 | Sequence 98, Appl |
| 4 | 140 | 100.0 | 42 | 3 US-09-100-414B-99 | Sequence 99, Appl |
| 5 | 140 | 100.0 | 42 | 3 US-09-100-414B-100 | Sequence 100, Appl |
| 6 | 140 | 100.0 | 42 | 4 US-09-303-323-98 | Sequence 98, Appl |
| 7 | 140 | 100.0 | 42 | 4 US-09-303-323-99 | Sequence 99, Appl |
| 8 | 140 | 100.0 | 42 | 4 US-09-303-323-100 | Sequence 100, Appl |
| 9 | 140 | 100.0 | 45 | 3 US-09-100-414B-101 | Sequence 101, Appl |
| 10 | 140 | 100.0 | 45 | 3 US-09-303-323-101 | Sequence 101, Appl |
| 11 | 140 | 100.0 | 46 | 3 US-09-100-414B-96 | Sequence 96, Appl |
| 12 | 140 | 100.0 | 46 | 3 US-09-100-414B-97 | Sequence 97, Appl |
| 13 | 140 | 100.0 | 46 | 4 US-09-303-323-96 | Sequence 96, Appl |
| 14 | 140 | 100.0 | 46 | 4 US-09-303-323-97 | Sequence 97, Appl |
| 15 | 140 | 100.0 | 63 | 3 US-09-100-414B-102 | Sequence 102, Appl |
| 16 | 140 | 100.0 | 63 | 4 US-09-303-323-102 | Sequence 102, Appl |
| 17 | 117 | 83.6 | 106 | 2 US-08-232-539D-54 | Sequence 54, Appl |
| 18 | 117 | 83.6 | 113 | 2 US-08-232-539D-56 | Sequence 56, Appl |
| 19 | 106.5 | 76.1 | 119 | 2 US-08-464-025A-1 | Sequence 1, Appl |
| 20 | 93 | 66.4 | 22 | 2 US-08-232-539D-19 | Sequence 19, Appl |
| 21 | 93 | 66.4 | 24 | 2 US-08-232-539D-20 | Sequence 20, Appl |
| 22 | 93 | 66.4 | 56 | 2 US-08-232-539D-18 | Sequence 18, Appl |
| 23 | 89 | 63.6 | 118 | 3 US-08-466-151-1 | Sequence 1, Appl |
| 24 | 85 | 60.7 | 426 | 1 US-08-336-583-2 | Sequence 2, Appl |
| 25 | 85 | 60.7 | 426 | 5 PCT-US95-13795-2 | Sequence 2, Appl |
| 26 | 49.5 | 35.4 | 1155 | 1 PCT-US96-948A-29 | Sequence 29, Appl |
| 27 | 49.5 | 35.4 | 1155 | 5 PCT-US96-09319-29 | Sequence 29, Appl |

| | | | | | | |
|----|------|------|------|---|-------------------|-------------------|
| 28 | 46.5 | 33.2 | 1234 | 2 | US-08-317-310A-15 | Sequence 15, Appl |
| 29 | 46.5 | 33.2 | 1234 | 5 | PCT-US95-13041-15 | Sequence 15, Appl |
| 30 | 46 | 32.9 | 50 | 1 | US-08-247-475-37 | Sequence 37, Appl |
| 31 | 46 | 32.9 | 50 | 1 | US-08-479-650-37 | Sequence 37, Appl |
| 32 | 46 | 32.9 | 50 | 1 | US-08-191-866D-59 | Sequence 59, Appl |
| 33 | 46 | 32.9 | 50 | 1 | US-08-674-169-37 | Sequence 37, Appl |
| 34 | 46 | 32.9 | 50 | 2 | US-08-185-949B-59 | Sequence 59, Appl |
| 35 | 46 | 32.9 | 91 | 2 | US-09-047-125-31 | Sequence 31, Appl |
| 36 | 46 | 32.9 | 91 | 3 | US-07-736-335E-31 | Sequence 31, Appl |
| 37 | 46 | 32.9 | 155 | 1 | US-08-150-203A-10 | Sequence 10, Appl |
| 38 | 46 | 32.9 | 155 | 1 | US-08-454-730-10 | Sequence 10, Appl |
| 39 | 45 | 32.1 | 109 | 2 | US-08-646-981-6 | Sequence 6, Appl |
| 40 | 44 | 31.4 | 338 | 3 | US-08-890-719-12 | Sequence 12, Appl |
| 41 | 44 | 31.4 | 355 | 3 | US-08-890-719-11 | Sequence 11, Appl |
| 42 | 44 | 31.4 | 355 | 3 | US-08-890-719-13 | Sequence 13, Appl |
| 43 | 44 | 31.4 | 483 | 4 | US-09-049-672A-5 | Sequence 5, Appl |
| 44 | 44 | 31.4 | 499 | 4 | US-09-049-672A-1 | Sequence 1, Appl |
| 45 | 44 | 31.4 | 1068 | 1 | US-08-396-479B-12 | Sequence 12, Appl |

ALIGNMENTS

RESULT 1
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-95

Query Match 100.0%; Score 140; DB 3; Length 25;
Best Local Similarity 100.0%; Pred. No. 1e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPRALMRSTTKC 25

Db 1 CGETYQSRVTHPLPRALMRSTTKC 25

RESULT 2
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-751-6849
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 100.0%; Score 140; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 1e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPRALMRSTTKC 25
|||||.....
DB 1 CGETYQSRVTHPLPRALMRSTTKC 25

RESULT 3
US-09-100-414B-98
; Sequence 98, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows

; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-751-6849
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-98

Query Match 100.0%; Score 140; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 2e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPRALMRSTTKC 25
|||||.....
DB 18 CGETYQSRVTHPLPRALMRSTTKC 42

RESULT 4
US-09-100-414B-99
; Sequence 99, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-751-6849
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-99

Query Match 100.0%; Score 140; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 2e-16;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
|||||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 5

US-09-100-414B-100
; Sequence 100, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-751-6849
; TELEFAX: 212-758-4800
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-100

Query Match 100.0%; Score 140; DB 3; Length 42;
Best Local Similarity 100.0%; Pred. No. 2e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
|||||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 6

US-09-303-323-98
; Sequence 98, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-751-6849
; TELEFAX: 212-758-4800
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-98

Query Match 100.0%; Score 140; DB 4; Length 42;
Best Local Similarity 100.0%; Pred. No. 2e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
|||||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 7

US-09-303-323-99
; Sequence 99, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-751-6849
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:

APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-101

Query Match 100.0%; Score 140; DB 4; Length 45;
Best Local Similarity 100.0%; Pred. No. 2.1e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
DB 21 CGETYQSRVTHPLPALMRSTTKC 45

RESULT 11
US-09-100-414B-96
Sequence 96, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-96

Query Match 100.0%; Score 140; DB 3; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.2e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25

DB 22 CGETYQSRVTHPLPALMRSTTKC 46
RESULT 12
US-09-100-414B-97
Sequence 97, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-97

Query Match 100.0%; Score 140; DB 3; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.2e-16;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
DB 22 CGETYQSRVTHPLPALMRSTTKC 46

RESULT 13
US-09-303-323-96
Sequence 96, Application US/09303323
Patent No. 6228987
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
TITLE OF INVENTION: IMMUNOGENS
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

Search completed: March 4, 2002, 12:56:34
Job time: 256 sec

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:58:23 ; Search time 80.02 Seconds
(without alignments)
23.142 Million cell updates/sec

Title: US-09-701-623C-5

Perfect score: 140

Sequence: 1 CGFTYQSRVTHPLPALMRSTTKC 25

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 220984

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 140 | 100.0 | 25 | 21 | AA79998 |
| 2 | 140 | 100.0 | 25 | 21 | Optimised IgE-CH3 |
| 3 | 140 | 100.0 | 25 | 21 | Modified human IgE |
| 4 | 108 | 77.1 | 25 | 21 | Peptide sequence o |
| 5 | 103 | 73.6 | 25 | 21 | Optimised IgE-CH3 |
| 6 | 94 | 67.1 | 25 | 21 | Optimised IgE-CH3 |
| 7 | 93 | 66.4 | 22 | 20 | Optimised IgE-CH3 |
| 8 | 93 | 66.4 | 24 | 20 | IgE peptide antago |
| 9 | 84 | 60.0 | 17 | 21 | IgE peptide antago |
| 10 | 80 | 57.1 | 17 | 21 | Antibody 15A.2 hum |
| 11 | 68 | 48.6 | 19 | 21 | Human IgE C epsilo |

| | | | | | | |
|----|------|------|----|----|---------|--------------------|
| 12 | 68 | 48.6 | 19 | 22 | AA51033 | IgE peptide #11. |
| 13 | 66 | 47.1 | 13 | 21 | AA26518 | Human IgE C epsilo |
| 14 | 64 | 45.7 | 20 | 18 | AA24102 | Canine immunoglobi |
| 15 | 64 | 45.7 | 25 | 21 | AA80001 | Optimised IgE-CH3 |
| 16 | 64 | 45.7 | 25 | 21 | AA80048 | IgE derived target |
| 17 | 60 | 42.9 | 17 | 21 | AA50893 | Antibody 15A.2 bin |
| 18 | 60 | 42.9 | 17 | 21 | AA50894 | Antibody 15A.2 can |
| 19 | 58.5 | 41.8 | 12 | 21 | AA26517 | Human IgE C epsilo |
| 20 | 55 | 39.3 | 14 | 21 | AA26519 | Human IgE C epsilo |
| 21 | 52 | 37.1 | 12 | 21 | AA26516 | Human IgE C epsilo |
| 22 | 52 | 37.1 | 12 | 22 | AA51038 | IgE peptide #16. |
| 23 | 50 | 35.7 | 17 | 21 | AA50897 | Antibody 15A.2 fel |
| 24 | 48 | 34.3 | 17 | 21 | AA50898 | Antibody 15A.2 swi |
| 25 | 41.5 | 29.6 | 16 | 21 | AA24095 | Human apoptosis re |
| 26 | 40 | 28.6 | 14 | 22 | AA00660 | Human protein frag |
| 27 | 39 | 27.9 | 21 | 22 | AA19280 | Peptide #5714 enco |
| 28 | 39 | 27.9 | 21 | 22 | AA32002 | Peptide #6039 enco |
| 29 | 39 | 27.9 | 25 | 13 | AA27321 | Peptide corresp. t |
| 30 | 37 | 26.4 | 9 | 21 | AA68599 | Monoclonal antibod |
| 31 | 37 | 26.4 | 9 | 21 | AA50891 | Antibody 15A.2 bin |
| 32 | 37 | 26.4 | 14 | 21 | AA69579 | Immunogenic peptid |
| 33 | 37 | 26.4 | 25 | 19 | AA21389 | Human HUPF-I mutan |
| 34 | 36 | 25.7 | 9 | 21 | AA69600 | Monoclonal antibod |
| 35 | 36 | 25.7 | 9 | 21 | AA50892 | Antibody 15A.2 bin |
| 36 | 35 | 25.0 | 16 | 21 | AA50900 | Antibody 15A.2 equ |
| 37 | 35 | 25.0 | 24 | 15 | AA53668 | HDag' antigen HF-2 |
| 38 | 35 | 25.0 | 25 | 15 | AA53659 | HDag' domain #1 of |
| 39 | 35 | 25.0 | 25 | 15 | AA53662 | HDag' domain #4 of |
| 40 | 35 | 25.0 | 25 | 15 | AA53665 | HDag' domain #7 of |
| 41 | 34 | 24.3 | 17 | 19 | AA21514 | Human neuroendocri |
| 42 | 34 | 24.3 | 25 | 15 | AA53661 | HDag' domain #3 of |
| 43 | 33 | 23.6 | 14 | 21 | AA81324 | Human growth hormo |
| 44 | 33 | 23.6 | 18 | 19 | AA66632 | HSV-2 glycoprotein |
| 45 | 33 | 23.6 | 21 | 17 | AA88855 | Peptide targetted |

ALIGNMENTS

RESULT 1

AA79998
ID AA79998 standard; Peptide; 25 AA.
XX
AC AA79998;
XX
DT 15-MAY-2000 (first entry)
XX
DE Optimised IgE-CH3 domain antigen peptide for human IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Homo sapiens.
OS Synthetic.

XX WO967293-A1.

XX 29-DEC-1999

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy

```

xx PS Claim 1; Page 21; 155pp; English.
xx CC The present invention describes immunoglobulin E (IgE)-CH3 domain
xx CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
xx CC and anti-asthmatic properties. (I) induces polyclonal antibodies
xx CC specific for a target effector site on the epsilon-heavy chain of IgE,
xx CC and so preventing triggering and activation of mast cells and basophils
xx CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
xx CC containing (I) are used for active immunisation against IgE-mediated
xx CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
xx CC dermatitis. Nucleic acids that encode these compounds are useful for
xx CC recombinant production of corresponding peptides or in DNA vaccines.
xx CC Conjugates of (I) that include a promiscuous T helper cell epitope
xx CC (functional in genetically diverse subjects), in addition to a B cell
xx CC target epitope, have increased immunogenicity and may include cyclic
xx CC constraints (disulfide bridge) to stabilise conformational features and
xx CC maximize cross-reactivity to the natural target. They induce safe
xx CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
xx CC acid sequences used in the exemplification of the present invention.
xx SQ Sequence 25 AA;

Query Match 100.0%; Score 140; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.8e-15;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPRALMRSTTKC 25
Db 1 cgetyqsrvtphlpralmrsttkc 25

RESULT 2
AAY91212
ID AAY91212 standard; peptide; 25 AA.
XX AC AAY91212;
XX DT 22-MAY-2000 (first entry)
XX DE Modified human IgE CH3 domain, SEQ ID NO:92.
XX KW Promiscuous T-cell epitope; measles virus F protein; MVF;
XX KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
XX KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
XX KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
XX KW foot and mouth disease virus; Immunoglobulin E; IgE; anti-allergic;
XX KW Plasmodium falciparum; circumsporozoite; antimalarial; CESTP;
XX KW cholesterol ester transport protein; anti-arteriosclerotic.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN WO9966957-A2.
XX PD 29-DEC-1999
XX PF 21-JUN-1999; 99WO-US13975.
XX PR 20-JUN-1998; 98US-0100412.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY;
XX DR WPI; 2000-160564/14.
XX PT New artificial T helper cell epitope and derived immunogens with target
XX PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
XX PT or human immune deficiency virus
XX PS Example 6; Page 40; 129pp; English.

```

```

xx CC The invention relates to novel promiscuous T helper cell epitopes (Th),
xx CC and immunogenic peptides comprising the Th epitopes of the invention
xx CC along with B cell epitopes. The Th epitopes and peptide immunogens
xx CC containing them, are used to induce a T helper cell response,
xx CC specifically against Plasmodium falciparum, cholesterol ester transport
xx CC protein (CESTP) or HIV epitopes, but more generally against any pathogen,
xx CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
xx CC peptide immunogens may be used for prevention and/or treatment of
xx CC infections (HIV, foot-and-mouth disease or malaria); for cancer
xx CC immunotherapy; for inhibition of the action of luteinising hormone
xx CC releasing hormone (LHRH) for contraception, treatment of hormone-
xx CC dependent cancer, prevention of boar taint in meat, and
xx CC immunocastration; for promoting the growth of animals; or for
xx CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
xx CC Th (functional in genetically diverse subjects) into an immunogen
xx CC improves capacity to induce a strong T helper cell-mediated immune
xx CC response, resulting in production of antibodies against a target
xx CC antigen. Th can replace carrier proteins and pathogen-derived T helper
xx CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
xx CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
xx CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
xx CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
xx CC from hepatitis B virus (HBV) surface antigen, and sequences
xx CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
xx CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
xx CC comprising a LHRH sequence joined to a promiscuous Th epitope. AAY91197
xx CC is the LHRH target antigenic peptide used in these LHRH antigenic
xx CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
xx CC peptides comprising somatostatin and a Th epitope. Somatostatin
xx CC immunogens may be used to promote growth in livestock. AAY91208 is a
xx CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are HIV Th
xx CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
xx CC infection of T cells. AAY90212 is a modified version of a human IgE
xx CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
xx CC antigenic peptides which may be used in the treatment of allergies.
xx CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
xx CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
xx CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
xx CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
xx CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
xx CC CESTP-derived peptides and AAY91232-Y91241 are immunogens comprising a
xx CC CESTP peptide and a Th epitope which may be used to prevent or treat
xx CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
xx CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
xx CC AAY91258-Y91273 are antigenic peptides comprising MVF Th and HIV-1 B-cell
xx CC epitope which may be used as a component in an anti-HIV-1 vaccine.
xx CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
xx CC protein epitope from Yersinia species, and hinge spacer peptide, both of
xx CC which may optionally be used in the antigenic peptides of the
xx CC invention.
xx SQ Sequence 25 AA;

```

Query Match 100.0%; Score 140; DB 21; Length 25;

Best Local Similarity 100.0%; Pred. No. 4.8e-15;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPRALMRSTTKC 25

Db 1 cgetyqsrvtphlpralmrsttkc 25

RESULT 3

AAY68602

ID AAY68602 standard; peptide; 25 AA.

XX AC AAY68602;

XX DT 05-MAY-2000 (first entry)

XX DE Peptide sequence of the invention.

XX KW Helper T cell epitope; peptide immunogen; LHRH;
KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
XX OS Unidentified.
XX PN WO9967293-A1.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-US13959.
XX PR 20-JUN-1998; 98US-0100287.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY; Walfield AM;
XX WPI; 2000-160578/14.
XX New peptide immunogen containing luteinising hormone-releasing hormone
XX antigen site and helper T cell epitope, for e.g. contraception and
XX treatment of cancer
XX Disclosure; Page 92; 102pp; English.
XX The specific helper T cell (Th) epitope and a target antigen, luteinising
XX hormone-releasing hormone (LHRH). The peptide immunogens cause
XX induction of a specific immune response to LHRH which is involved in
XX regulation of spermatogenesis, ovulation, oestrus, sexual development
XX and secretion of sex hormones. Provision of a promiscuous T helper
XX epitope (which is functional in genetically diverse subjects) provides
XX optimum immunogenicity to the B cell epitopes of the target antigen and
XX thus high antibody titres against the target antigen. The peptide
XX immunogens of the invention are used to vaccinate against mammalian LHRH,
XX for use as (reversible) contraceptive; control of hormone-dependent
XX tumours (cancer of prostate or breast, also endometriosis); to prevent
XX boar taint (and improve meat quality) and for immunocastration. The
XX present sequence appears in the specification.
XX SQ Sequence 25 AA;
Query Match 100.0%; Score 140; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.8e-15;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGETYQSRVTHPLPALMRSTTKC 25
Db 1 cgetysrvthplpalmrsttkc 25
RESULT 4
AAY79999
ID AAY79999 standard; Peptide; 25 AA.
AC AAY79999;
XX
DT 15-MAY-2000 (first entry)
DE Optimised IgE-CH3 domain antigen peptide for dog IgE.
XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
OS Canis sp.
OS Synthetic.

XX PN WO9967293-A1.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-US13959.
XX PR 20-JUN-1998; 98US-0100287.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY; Walfield AM;
XX WPI; 2000-160578/14.
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX for immunization against allergy
XX Claim 1; Page 99; 155pp; English.
XX The present invention describes immunoglobulin E (IgE)-CH3 domain
XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX and anti-asthmatic properties. (I) induces polyclonal antibodies
XX specific for a target effector site on the epsilon-heavy chain of IgE,
XX and so preventing triggering and activation of mast cells and basophils
XX and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX containing (I) are used for active immunisation against IgE-mediated
XX allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX dermatitis. Nucleic acids that encode these compounds are useful for
XX recombinant production of corresponding peptides or in DNA vaccines.
XX Conjugates of (I) that include a promiscuous T helper cell epitope
XX (functional in genetically diverse subjects), in addition to a B cell
XX target epitope, have increased immunogenicity and may include cyclic
XX constraints (disulfide bridge) to stabilise conformational features and
XX maximize cross-reactivity to the natural target. They induce safe
XX (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
XX acid sequences used in the exemplification of the present invention.
XX SQ Sequence 25 AA;
Query Match 77.1%; Score 108; DB 21; Length 25;
Best Local Similarity 72.0%; Pred. No. 4.6e-10;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 CGETYQSRVTHPLPALMRSTTKC 25
Db 1 cgetysrvthplpalmrsttkc 25
RESULT 5
AAY80000
ID AAY80000 standard; Peptide; 25 AA.
XX
AC AAY80000;
XX
DT 15-MAY-2000 (first entry)
DE Optimised IgE-CH3 domain antigen peptide for rat IgE.
XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
OS Rattus sp.
OS Synthetic.
XX
PN WO9967293-A1.
XX PD 29-DEC-1999.
XX PR 21-JUN-1999; 99WO-US13959.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY; Walfield AM;
XX WPI; 2000-160578/14.
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX for immunization against allergy
XX Claim 1; Page 99; 155pp; English.
XX The present invention describes immunoglobulin E (IgE)-CH3 domain
XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX and anti-asthmatic properties. (I) induces polyclonal antibodies
XX specific for a target effector site on the epsilon-heavy chain of IgE,
XX and so preventing triggering and activation of mast cells and basophils
XX and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX containing (I) are used for active immunisation against IgE-mediated
XX allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX dermatitis. Nucleic acids that encode these compounds are useful for
XX recombinant production of corresponding peptides or in DNA vaccines.
XX Conjugates of (I) that include a promiscuous T helper cell epitope
XX (functional in genetically diverse subjects), in addition to a B cell
XX target epitope, have increased immunogenicity and may include cyclic
XX constraints (disulfide bridge) to stabilise conformational features and
XX maximize cross-reactivity to the natural target. They induce safe
XX (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
XX acid sequences used in the exemplification of the present invention.
XX SQ Sequence 25 AA;

XX PR 20-JUN-1998; 98US-0100287.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY, Walfield AM;
XX DR WPI; 2000-160578/14.
XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy -
XX PS Claim 1; Page 99; 155pp; English.
XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.
XX SQ Sequence 25 AA;

Query Match 73.6%; Score 103; DB 21; Length 25;
Best Local Similarity 68.0%; Pred. No. 2.8e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
Db 1 cgegyqsrvdhphkpkpivrsitkc 25
||| ||||| ||| | :|| |||
1 cgegyqsrvdhphkpkpivrsitkc 25

RESULT 6
AAY80077
ID AAY80077 standard; Peptide; 25 AA.
XX AC AAY80077;
XX DT 15-MAY-2000 (first entry)
XX DE Optimised IgE-CH3 domain antigen peptide for horse IgE.
XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX OS Equus caballus.
OS Synthetic.
XX PN WO9967293-A1.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-US13959.
XX PR 20-JUN-1998; 98US-0100287.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY, Walfield AM;

XX DR WPI; 2000-160578/14.
XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy -
XX PS Claim 1; Page 146; 155pp; English.
XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.
XX SQ Sequence 25 AA;

Query Match 67.1%; Score 94; DB 21; Length 25;
Best Local Similarity 64.0%; Pred. No. 7e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPALMRSTTKC 25
Db 1 cgetyqstsvahpdlprevvrsiack 25
||||| | :|| ||| :|| ||
1 cgetyqstsvahpdlprevvrsiack 25

RESULT 7
AAY42585
ID AAY42585 standard; peptide; 22 AA.
XX AC AAY42585;
XX DT 10-JAN-2000 (first entry)
XX DE IgE peptide antagonist.
XX KW Immunoglobulin E; IgE; antagonist; FcepsilonRI receptor; human; bds;
KW receptor-binding; binding determinant sequence; anti-IgE antibody;
KW allergic disease.
XX OS Homo sapiens.
XX PN US5965709-A.
XX PD 12-OCT-1999.
XX PF 21-APR-1994; 94US-0232539.
XX PR 14-AUG-1991; 91US-0744768.
PR 07-JAN-1994; 94US-0178583.
XX PA (GETH) GENENTECH INC.
XX PI Jardieu PM, Presta LG;
XX DR WPI; 1999-579941/49.
XX PT Immunoglobulin E variants as peptide antagonists useful for raising and
PT screening anti-immunoglobulin E (IgE) antibodies, in the isolation and
PT purification of FcepsilonRI receptor and in the treatment of allergic
PT diseases -

XX Disclosure; Column 9; 37pp; English.

XX The invention provides immunoglobulin E (IgE) antagonists comprising one

CC or more of the Fcεpsilon1RI receptor-binding determinant sites of human

CC IgE. The antagonists include IgE variants comprising an immunoglobulin

CC template and binding determinant sequences (bds) CDbds, EFBds and the

CC sequence shown in AAY42581. The CDbds (CD loop binding determinant

CC sequence) are selected from the sequences shown in AAY42567-Y42577 and

CC the EFBds (EF loop binding determinant sequence) are selected from

CC sequences shown in AAY42578-Y42580. The variants are useful in raising

CC and screening anti-IgE antibodies, in the isolation and purification of

CC Fcεpsilon1RI receptor and in the treatment and prophylaxis of allergic

XX diseases.

XX Sequence 22 AA;

Query Match 66.4%; Score 93; DB 20; Length 22;

Best Local Similarity 94.4%; Pred. No. 8.6e-08;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHLPALM 19

Db 5 getyqcrvthphlpralm 22

||||| |||||||||

RESULT 8

AAV42586

ID AAY42586 standard; peptide; 24 AA.

XX AC AAY42586;

XX 10-JAN-2000 (first entry)

XX DE IgE peptide antagonist.

XX Immunoglobulin E; IgE; antagonist; Fcεpsilon1RI receptor; human; bds;

KW receptor-binding; binding determinant sequence; anti-IgE antibody;

KX allergic disease.

XX Homo sapiens.

XX US5965709-A.

XX 12-OCT-1999.

XX 21-APR-1994; 94US-0232539.

XX 14-AUG-1991; 91US-0744768.

PR 07-JAN-1994; 94US-0178583.

XX (GETH) GENENTECH INC.

XX PA Jardieu PM, Presta LG;

XX PI WPI; 1999-579941/49.

XX Immunoglobulin E variants as peptide antagonists useful for raising and

PT screening anti-immunoglobulin E (IgE) antibodies, in the isolation and

PT purification of Fcεpsilon1RI receptor and in the treatment of allergic

PT diseases -

XX Disclosure; Column 9; 37pp; English.

XX The invention provides immunoglobulin E (IgE) antagonists comprising one

CC or more of the Fcεpsilon1RI receptor-binding determinant sites of human

CC IgE. The antagonists include IgE variants comprising an immunoglobulin

CC template and binding determinant sequences (bds) CDbds, EFBds and the

CC sequence shown in AAY42581. The CDbds (CD loop binding determinant

CC sequence) are selected from the sequences shown in AAY42567-Y42577 and

CC the EFBds (EF loop binding determinant sequence) are selected from

CC sequences shown in AAY42578-Y42580. The variants are useful in raising

CC and screening anti-IgE antibodies, in the isolation and purification of

CC Fcεpsilon1RI receptor and in the treatment and prophylaxis of allergic

CC diseases.

XX Sequence 24 AA;

Query Match 66.4%; Score 93; DB 20; Length 24;

Best Local Similarity 94.4%; Pred. No. 9.5e-08;

Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHLPALM 19

Db 7 getyqcrvthphlpralm 24

||||| |||||||||

RESULT 9

AAV50895

ID AAY50895 standard; peptide; 17 AA.

XX AC AAY50895;

XX 24-FEB-2000 (first entry)

XX DE Antibody 15A.2 human IgE binding epitope 1.

XX Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;

KW epitope; prophylaxis; treatment; mimotope.

XX OS Synthetic.

XX EP957111-A2.

XX 17-NOV-1999.

XX 09-APR-1999; 99EP-0107035.

XX 09-APR-1998; 98US-0058331.

PR 30-MAR-1999; 99US-0281760.

XX (IDEX-) IDEXX LAB INC.

XX Lawton R, Mermer B, Francoeur G;

XX WPI; 2000-040833/04.

XX Binding proteins used for treatment or prophylaxis of canine allergy -

XX Disclosure; Fig 7; 30pp; English.

XX This invention describes a novel binding protein which specifically

CC binds to native canine free or B-cell bound IgE, and which doesn't bind

CC to IgE when the IgE is bound to mast cells. The peptide products of the

CC invention have anti-allergic activity. The antibodies bind to defined

CC epitopes on free or B-cell bound IgE molecules which have an important

CC role in allergic reaction. The specific binding proteins are used to

CC produce a pharmaceutical composition, preferably with a diluent, which

CC can be used for prophylaxis or treatment of canine allergy.

CC AAY50876-Y50900 represent peptide mimotopes used in the method of the

CC invention.

XX Sequence 17 AA;

Query Match 60.0%; Score 84; DB 21; Length 17;

Best Local Similarity 94.1%; Pred. No. 1.6e-06;

Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 QSRVTHPHLPALMRST 22

Db 1 qcrvthphlpralmrst 17

||||| |||||||||

XX Prieels J;
XX WPI; 2001-091150/10.
XX New vaccine comprising allergy peptides linked by an inert carrier,
PT useful for boosting an anti-allergy immune response in an individual
PT susceptible to an allergic response
XX Claim 5; Page 21; 26pp; English.
XX The present invention relates to a composition comprising allergy
CC peptides linked by an inert carrier. The allergy peptides are derived
CC from immunoglobulin E (IgE) or IgE receptor. The present peptide is one
CC such peptide from IgE. The composition is useful as a vaccine or for
CC manufacturing a medicament for the prophylaxis or treatment of allergy.
CC In particular, for boosting an anti-allergy immune response in an
CC individual susceptible to an allergic response.
XX Sequence 19 AA;
SQ

Query Match 48.6%; Score 68; DB 22; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PHLPRALMRSTTK 24
| | | | | | | | | |
DB 1 phlpralmrsttk 13

RESULT 13
AAB26518
ID AAB26518 standard; Peptide; 13 AA.
XX
AC AAB26518;
XX
DT 11-JAN-2001 (first entry)
XX
DE Human IgE C epsilon3/4 domain mimotope P13b.
XX
KW IgE; C epsilon3; C epsilon4; histamine release inhibitor; vaccine;
KW antibody; epitope; mimotope; human.
XX
OS Homo sapiens.
OS Synthetic.

XX Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT
XX
XX WO200050461-A1.
XX
XX 31-AUG-2000.
XX
XX 22-FEB-2000; 2000WO-EP01456.
XX
XX 25-FEB-1999; 99GB-0004408.
XX
XX 21-JUL-1999; 99GB-0017144.
XX
XX 07-AUG-1999; 99GB-0018598.
XX
XX 07-AUG-1999; 99GB-0018599.
XX
XX 07-AUG-1999; 99GB-0018601.
XX
XX 07-AUG-1999; 99GB-0018604.
XX
XX 07-AUG-1999; 99GB-0018606.
XX
XX 29-OCT-1999; 99GB-0025618.

XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA (PEPT-) PEPTIDE THERAPEUTICS LTD.
XX

XX Friede M, Mason S, Turnell WG, Van Mechelen WP;
PI Vinals De Bassols YC;
XX WPI; 2000-572074/53.
DR

XX Peptides comprising surface exposed epitopes or mimotopes derived from
PT C-epsilon3 or C-epsilon4 domains of IgE, useful for preventing or
PT treating allergy
XX
XX Disclosure; Page 8; 76pp; English.
XX
XX The present invention relates epitopes and mimotopes of an isolated
CC surface exposed epitope of C epsilon3 or C epsilon4 domain of IgE. The
CC epitopes were identified by calculating the accessible surface of each
CC IgE residue. Mimotopes were designed to be similar to the epitopes.
CC The epitopes are useful in preparing medicaments for treating or
CC preventing allergies. The epitopes and mimotopes of the invention
CC induce anti-IgE antibodies which are capable of raising
CC non-anaphylactic antibodies and inhibiting histamine release. The
CC present sequence is an IgE C epsilon3/4 domain mimotope.
XX Sequence 13 AA;
SQ

Query Match 47.1%; Score 66; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00076;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 THPLPRALMRS 21
| | | | | | | | | |
DB 1 thplpralrms 12

RESULT 14
AAW24102
ID AAW24102 standard; peptide; 20 AA.
XX
AC AAW24102;
XX
DT 21-NOV-1997 (first entry)
XX
DE Canine immunoglobulin E peptide 5.
XX
KW Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.
XX
OS Canis familiaris.
XX
XX JF09169795-A.
XX
XX 30-JUN-1997.
XX
XX 22-DEC-1995; 95JP-0334381.
XX
XX 22-DEC-1995; 95JP-0334381.
XX
XX (HITB) HITACHI CHEM CO LTD.
XX
XX WPI; 1997-389423/36.
XX
XX N-PSDB; AAT85651.
XX
XX Canine immunoglobulin E peptide fragment and related DNA - useful
PT for the preparation of anti-canine immunoglobulin E antibody
XX
XX Claim 2; Page 9; 12pp; Japanese.
XX

XX AAW24098-106 are peptide fragments containing at least 5 continuous
CC amino acids of the partial canine immunoglobulin E (IgE) protein shown
CC in AAW24097. The peptides are used for the preparation of anti-canine
CC IgE antibody. The anti-canine IgE antibody can be used for the diagnosis
CC of canine allergies.
XX
XX Sequence 20 AA;
SQ

Query Match 45.7%; Score 64; DB 18; Length 20;
Best Local Similarity 64.7%; Pred. No. 0.0025;
Matches 11; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 8 RVTHPLRALMRSTTK 24
|||||: :|||
Db 3 rvthplpkdivrsiak 19

RESULT 15

AAAY80001
ID AAY80001 standard; Peptide; 25 AA.

AC AAY80001;

DT 15-MAY-2000 (first entry)

XX Optimised IgE-CH3 domain antigen peptide for mouse IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Mus sp.

OS Synthetic.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX for immunization against allergy

PS Claim 1; Page 100; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX and anti-asthmatic properties. (I) induces polyclonal antibodies
XX specific for a target effector site on the epsilon-heavy chain of IgE,
XX and so preventing triggering and activation of mast cells and basophils
XX and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX containing (I) are used for active immunisation against IgE-mediated
XX allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX dermatitis. Nucleic acids that encode these compounds are useful for
XX recombinant production of corresponding peptides or in DNA vaccines.
XX Conjugates of (I) that include a promiscuous T helper cell epitope
XX (functional in genetically diverse subjects), in addition to a B cell
XX target epitope, have increased immunogenicity and may include cyclic
XX constraints (disulfide bridge) to stabilise conformational features and
XX maximize cross-reactivity to the natural target. They induce safe
XX (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
XX acid sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

Query Match 45.7%; Score 64; DB 21; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0033;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CGETYQSRVTHPLRALMRSTTKC 25
||| | | | : ||| | |
Db 1 cgyyqgsivdrpdpfpkpvrsitic 25

Search completed: March 4, 2002, 13:10:08
Job time: 705 sec

10/10/01 10:10:10 AM

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:10:09 ; Search time 80.02 Seconds
(without alignments)
23.142 Million cell updates/sec

Title: US-09-701-623C-84
Perfect score: 135
Sequence: 1 GGEYKSTVSHPDLPREVRSIAKC 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 220984

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 135 | 100.0 | 25 | 21 | AAV80077 |
| 2 | 105 | 77.8 | 25 | 21 | AAV79999 |
| 3 | 94 | 69.6 | 25 | 21 | AAV79998 |
| 4 | 94 | 69.6 | 25 | 21 | AAV91212 |
| 5 | 94 | 69.6 | 25 | 21 | AAV68602 |
| 6 | 89 | 65.9 | 25 | 21 | AAV80000 |
| 7 | 74 | 54.8 | 25 | 21 | AAV80001 |
| 8 | 63 | 46.7 | 20 | 18 | AAW24102 |
| 9 | 59 | 43.7 | 16 | 21 | AAV50900 |
| 10 | 58 | 43.0 | 22 | 20 | AAV42585 |
| 11 | 58 | 43.0 | 24 | 20 | AAV42586 |
| | | | | | Optimised IgE-CH3 |
| | | | | | Optimised IgE-CH3 |
| | | | | | Optimised IgE-CH3 |
| | | | | | Modified human IgE |
| | | | | | Peptide sequence o |
| | | | | | Optimised IgE-CH3 |
| | | | | | Optimised IgE-CH3 |
| | | | | | Canine immunoglobi |
| | | | | | Antibody 15A.2 equ |
| | | | | | IgE peptide antago |
| | | | | | IgE peptide antago |

12 55 40.7 17 21 AAV50898
13 54 40.0 17 21 AAV50893
14 54 40.0 17 21 AAV50894
15 53 39.3 17 21 AAV50897
16 48 35.6 17 21 AAV50896
17 45 33.3 17 21 AAV50895
18 42 31.1 25 21 AAV80048
19 41 30.4 13 21 AAB26518
20 41 30.4 15 21 AAV50899
21 41 30.4 22 16 AAR72686
22 39 28.9 22 16 AAR72687
23 36 26.7 19 21 AAB26503
24 36 26.7 19 22 AAB51033
25 35 25.9 15 20 AAV41939
26 35 25.9 18 19 AAW66628
27 35 25.9 23 22 AAM36849
28 35 25.9 24 20 AAV25671
29 34 25.2 21 21 AAV58800
30 34 25.2 21 22 AAG77526
31 34 25.2 23 22 AAB70581
32 34 25.2 25 20 AAY39193
33 34 25.2 25 20 AAY39050
34 33.5 24.8 12 21 AAB26517
35 33 24.4 16 21 AAV59357
36 33 24.4 17 11 AAR02179
37 33 24.4 20 18 AAW24103
38 32.5 24.1 21 16 AAR74085
39 32 23.7 14 21 AAB26519
40 32 23.7 15 19 AAW85191
41 31 23.0 12 21 AAB26516
42 31 23.0 12 22 AAB51038
43 31 23.0 15 21 AAV51953
44 31 23.0 15 21 AAV51965
45 31 23.0 16 18 AAW18594

ALIGNMENTS

RESULT 1
AAV80077
ID AAV80077 standard; Peptide; 25 AA.

XX AAV80077;

DT 15-MAY-2000 (first entry)

DE Optimised IgE-CH3 domain antigen peptide for horse IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Equus caballus.
OS Synthetic.

XX WO9967293-A1.

PD 29-DEC-1999

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy

CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY7994 to AAY8004 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;

Query Match 69.6%; Score 94; DB 21; Length 25;
 Best Local Similarity 64.0%; Pred. No. 4.1e-08;
 Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYSKTVSHPDLPREVRSIAKC 25
 Db 1 cgetysrvthphlpralmrsttkc 25

RESULT 4

AY91212
 ID AAY91212 standard; peptide; 25 AA.

XX
 AC AAY91212;

XX
 DT 22-MAY-2000 (first entry)

XX
 DE Modified human IgE CH3 domain, SEQ ID NO:92.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumporoite; antimalarial; CERP;
 KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX Homo sapiens.

OS Synthetic.

XX WO9666957-A2.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13975.

XX 20-JUN-1998; 98US-0100412.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI; 2000-160564/14.

XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus

XX Example 6; Page 40; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesteryl ester transport
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-

CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration); for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumporoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.
 XX
 SQ Sequence 25 AA;

Query Match 69.6%; Score 94; DB 21; Length 25;
 Best Local Similarity 64.0%; Pred. No. 4.1e-08;
 Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYSKTVSHPDLPREVRSIAKC 25
 Db 1 cgetysrvthphlpralmrsttkc 25

RESULT 5

AY68602

ID AAY68602 standard; peptide; 25 AA.

XX
 AC AAY68602;

XX 05-MAY-2000 (first entry)

XX Peptide sequence of the invention.

XX Helper T cell epitope; peptide immunogen; LHRH;
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
 XX Unidentified.

XX WO9666952-A1.

XX 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13960.
XX PR 20-JUN-1998; 98US-0100414.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY;
XX DR WPI; 2000-160578/14.
XX DR New peptide immunogen containing luteinising hormone-releasing hormone
XX PT antigen site and helper T cell epitope, for e.g. contraception and
XX PT treatment of cancer
XX PS Disclosure; Page 92; 102pp; English.
XX CC The specification describes peptide immunogens comprising a
CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
CC hormone-releasing hormone (LHRH). The peptide immunogens cause
CC induction of a specific immune response to LHRH which is involved in
CC regulation of spermatogenesis, ovulation, oestrus, sexual development
CC and secretion of sex hormones. Provision of a promiscuous T helper
CC epitope (which is functional in genetically diverse subjects) provides
CC optimum immunogenicity to the B cell epitopes of the target antigen and
CC thus high antibody titres against the target antigen. The peptide
CC immunogens of the invention are used to vaccinate against mammalian LHRH,
CC for use as (reversible) contraceptive; control of hormone-dependent
CC tumours (cancer of prostate or breast; also endometriosis); to prevent
CC boar taint (and improve meat quality) and for immunocastration. The
CC present sequence appears in the specification.
XX CC
XX SQ Sequence 25 AA;

Query Match 69.6%; Score 94; DB 21; Length 25;
Best Local Similarity 64.0%; Pred. No. 4.le-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 1 CGETYKSTVSHPDLPREVRSIAKC 25
DB 1 cgetyqsrvtchphlpraimrsttkc 25

RESULT 6
AAY80000
ID AAY80000 standard; Peptide; 25 AA.
XX AC AAY80000;
XX DT 15-MAY-2000 (first entry)
XX DE Optimised IgE-CH3 domain antigen peptide for rat IgE.
XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX OS Rattus sp.
XX OS Synthetic.
XX PN WO9967293-A1.
XX PD 29-DEC-1999.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PF 21-JUN-1999; 99WO-US13959.
XX PR 20-JUN-1998; 98US-0100287.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY, Walfield AM;

XX DR WPI; 2000-160578/14.
XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX PS for immunization against allergy
XX PS Claim 1; Page 99; 155pp; English.
XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies
XX CC specific for a target effector site on the epsilon-heavy chain of IgE,
XX CC and so preventing triggering and activation of mast cells and basophils
XX CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX CC containing (I) are used for active immunisation against IgE-mediated
XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX CC dermatitis. Nucleic acids that encode these compounds are useful for
XX CC recombinant production of corresponding peptides or in DNA vaccines.
XX CC Conjugates of (I) that include a promiscuous T helper cell epitope
XX CC (functional in genetically diverse subjects), in addition to a B cell
XX CC target epitope, have increased immunogenicity and may include cyclic
XX CC constraints (disulfide bridge) to stabilise conformational features and
XX CC maximize cross-reactivity to the natural target. They induce safe
XX CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
XX CC acid sequences used in the exemplification of the present invention.
XX SQ Sequence 25 AA;

Query Match 65.9%; Score 89; DB 21; Length 25;
Best Local Similarity 60.0%; Pred. No. 2.5e-07;
Matches 15; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHPDLPREVRSIAKC 25
DB 1 cgegygsrvdhphfpkpvirsitkc 25

RESULT 7
AAY80001
ID AAY80001 standard; Peptide; 25 AA.
XX AC AAY80001;
XX DT 15-MAY-2000 (first entry)
XX DE Optimised IgE-CH3 domain antigen peptide for mouse IgE.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX OS Mus sp.
XX OS Synthetic.
XX PN WO9967293-A1.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-US13959.
XX PR 20-JUN-1998; 98US-0100287.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY, Walfield AM;
XX DR WPI; 2000-160578/14.
XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX PT for immunization against allergy

PS Claim 1; Page 100; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE.
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 25 AA;

Query Match 54.8%; Score 74; DB 21; Length 25;
 Best Local Similarity 52.0%; Pred. No. 6.2e-05;
 Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHPDLPREVRSIAKC 25
 II I I I I I I I I I I
 DB 1 cgyyqgsivdrpdkpivrsitlc 25

RESULT 8
 AAW24102
 ID AAW24102 standard; peptide; 20 AA.
 XX AC AAW24102;
 XX DT 21-NOV-1997 (first entry)
 XX DE Canine Immunoglobulin E peptide 5.
 XX KW Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.
 XX OS Canis familiaris.
 XX PN JP09169795-A.
 XX PD 30-JUN-1997.
 XX PF 22-DEC-1995; 95JP-0334381.
 XX PR 22-DEC-1995; 95JP-0334381.
 XX PS (HITB) HITACHI CHEM CO LTD.
 XX WPI; 1997-389423/36.
 DR N-PSDB; AAT85651.
 XX PT Canine immunoglobulin E peptide fragment and related DNA - useful
 PT for the preparation of anti-canine immunoglobulin E antibody
 XX PS Claim 2; Page 9; 12pp; Japanese.
 XX CC AAW24098-106 are peptide fragments containing at least 5 continuous
 CC amino acids of the partial canine immunoglobulin E (IgE) protein shown
 CC in AAW24097. The peptides are used for the preparation of anti-canine
 CC IgE antibody. The anti-canine IgE antibody can be used for the diagnosis
 CC of canine allergies.

XX SQ Sequence 20 AA;

Query Match 46.7%; Score 63; DB 18; Length 20;
 Best Local Similarity 68.8%; Pred. No. 0.0026;
 Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 9 VSHPDLPREVRSIAK 24
 I I I I I I I I I I I I
 DB 4 vtphlpkdivrsiak 19

RESULT 9
 AAY50900
 ID AAY50900 standard; peptide; 16 AA.
 XX AC AAY50900;
 XX DT 24-FEB-2000 (first entry)
 XX DE Antibody 15A.2 equine IgE binding epitope 1.
 XX KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 KW epitope; prophylaxis; treatment; mimotope.
 XX OS Synthetic.
 XX PN EP957111-A2.
 XX PD 17-NOV-1999.
 XX PF 09-APR-1999; 99EP-0107035.
 XX PR 09-APR-1998; 98US-0058331.
 XX PR 30-MAR-1999; 99US-0281760.
 XX PA (IDEX-) IDEXX LAB INC.
 XX PI Lawton R, Mermer B, Francoeur G;
 XX WPI; 2000-040833/04.
 XX PT Binding proteins used for treatment or prophylaxis of canine allergy -
 XX PS Disclosure; Fig 7; 30pp; English.
 XX CC This invention describes a novel binding protein which specifically
 CC binds to native canine free or B-cell bound IgE, and which doesn't bind
 CC to IgE when the IgE is bound to mast cells. The peptide products of the
 CC invention have anti-allergic activity. The antibodies bind to defined
 CC epitopes on free or B-cell bound IgE molecules which have an important
 CC role in allergic reaction. The specific binding proteins are used to
 CC produce a pharmaceutical composition, preferably with a diluent, which
 CC can be used for prophylaxis or treatment of canine allergy.
 CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.

XX SQ Sequence 16 AA;

Query Match 43.7%; Score 59; DB 21; Length 16;
 Best Local Similarity 91.7%; Pred. No. 0.0087;
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 KSTVSHPDLPRE 17
 I I I I I I I I I I I I
 DB 1 kctvshpdlpre 12

RESULT 10
 AAY42585
 ID AAY42585 standard; peptide; 22 AA.
 XX AC AAY42585;
 XX DT 10-JAN-2000 (first entry)

```
XX DE IgE peptide antagonist.
XX KW Immunoglobulin E; IgE; antagonist; FcepsilonRI receptor; human; bds;
XX KW receptor-binding; binding determinant sequence; anti-IgE antibody;
XX KW allergic disease.
XX OS Homo sapiens.
XX PN US5965709-A.
XX PD 12-OCT-1999.
XX PF 21-APR-1994; 94US-0232539.
XX PR 14-AUG-1991; 91US-0744768.
XX PR 07-JAN-1994; 94US-0178583.
XX PA (GETH ) GENENTECH INC.
XX PI Jardieu PM, Presta LG;
XX PI WPI; 1999-579941/49.
XX PT Immunoglobulin E variants as peptide antagonists useful for raising and
XX PT screening anti-immunoglobulin E (IgE) antibodies, in the isolation and
XX PT purification of FcepsilonRI receptor and in the treatment of allergic
XX PT diseases -
XX PS Disclosure; Column 9; 37pp; English.
XX CC The invention provides immunoglobulin E (IgE) antagonists comprising one
XX CC or more of the FcepsilonRI receptor-binding determinant sites of human
XX CC IgE. The antagonists include IgE variants comprising an immunoglobulin
XX CC template and binding determinant sequences (bds) CDBds, EFBds and the
XX CC sequence shown in AAY42581. The CDBds (CD loop binding determinant
XX CC sequence) are selected from the sequences shown in AAY42567-Y42577 and
XX CC the EFBds (EF loop binding determinant sequence) are selected from
XX CC sequences shown in AAY42578-Y42580. The variants are useful in raising
XX CC and screening anti-IgE antibodies, in the isolation and purification of
XX CC FcepsilonRI receptor and in the treatment and prophylaxis of allergic
XX CC diseases.
XX SQ Sequence 22 AA;

Query Match 43.0%; Score 58; DB 20; Length 22;
Best Local Similarity 55.6%; Pred. No. 0.018;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREV 19
   ||||| 1:11 ||| ::
DB 5 getyqcrvthphlpralm 22

RESULT 11
AAY42586
ID AAY42586 standard; peptide; 24 AA.
XX AC AAY42586;
XX DT 10-JAN-2000 (first entry)
XX DE IgE peptide antagonist.
XX KW Immunoglobulin E; IgE; antagonist; FcepsilonRI receptor; human; bds;
XX KW receptor-binding; binding determinant sequence; anti-IgE antibody;
XX KW allergic disease.
XX OS Homo sapiens.
XX PN US5965709-A.
XX PD 12-OCT-1999.
XX PF 21-APR-1994; 94US-0232539.
XX PR 14-AUG-1991; 91US-0744768.
XX PR 07-JAN-1994; 94US-0178583.
XX PA (GETH ) GENENTECH INC.
XX PI Jardieu PM, Presta LG;
XX PI WPI; 1999-579941/49.
XX PT Immunoglobulin E variants as peptide antagonists useful for raising and
XX PT screening anti-immunoglobulin E (IgE) antibodies, in the isolation and
XX PT purification of FcepsilonRI receptor and in the treatment of allergic
XX PT diseases -
XX PS Disclosure; Column 9; 37pp; English.
XX CC The invention provides immunoglobulin E (IgE) antagonists comprising one
XX CC or more of the FcepsilonRI receptor-binding determinant sites of human
XX CC IgE. The antagonists include IgE variants comprising an immunoglobulin
XX CC template and binding determinant sequences (bds) CDBds, EFBds and the
XX CC sequence shown in AAY42581. The CDBds (CD loop binding determinant
XX CC sequence) are selected from the sequences shown in AAY42567-Y42577 and
XX CC the EFBds (EF loop binding determinant sequence) are selected from
XX CC sequences shown in AAY42578-Y42580. The variants are useful in raising
XX CC and screening anti-IgE antibodies, in the isolation and purification of
XX CC FcepsilonRI receptor and in the treatment and prophylaxis of allergic
XX CC diseases.
XX SQ Sequence 22 AA;

Query Match 43.0%; Score 58; DB 20; Length 22;
Best Local Similarity 55.6%; Pred. No. 0.018;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREV 19
   ||||| 1:11 ||| ::
DB 5 getyqcrvthphlpralm 22

RESULT 12
AAY50898
ID AAY50898 standard; peptide; 17 AA.
XX AC AAY50898;
XX DT 24-FEB-2000 (first entry)
XX DE Antibody 15A.2 swine IgE binding epitope 1.
XX KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
XX KW epitope; prophylaxis; treatment; mimotope.
XX OS Synthetic.
XX PN EP957111-A2.
XX PD 17-NOV-1999.
XX PF 09-APR-1999; 99EP-0107035.
XX PR 09-APR-1998; 98US-0058331.
XX PR 30-MAR-1999; 99US-0281760.
XX PA (IDEX-) IDEXX LAB INC.
XX PI Lawton R, Mermer B, Francoeur G;
XX PD WPI; 2000-040833/04.
```

```
PD 12-OCT-1999.
XX 21-APR-1994; 94US-0232539.
XX 14-AUG-1991; 91US-0744768.
XX 07-JAN-1994; 94US-0178583.
XX (GETH ) GENENTECH INC.
XX Jardieu PM, Presta LG;
XX WPI; 1999-579941/49.
XX Immunoglobulin E variants as peptide antagonists useful for raising and
XX screening anti-immunoglobulin E (IgE) antibodies, in the isolation and
XX purification of FcepsilonRI receptor and in the treatment of allergic
XX diseases -
XX Disclosure; Column 9; 37pp; English.
XX The invention provides immunoglobulin E (IgE) antagonists comprising one
XX or more of the FcepsilonRI receptor-binding determinant sites of human
XX IgE. The antagonists include IgE variants comprising an immunoglobulin
XX template and binding determinant sequences (bds) CDBds, EFBds and the
XX sequence shown in AAY42581. The CDBds (CD loop binding determinant
XX sequence) are selected from the sequences shown in AAY42567-Y42577 and
XX the EFBds (EF loop binding determinant sequence) are selected from
XX sequences shown in AAY42578-Y42580. The variants are useful in raising
XX and screening anti-IgE antibodies, in the isolation and purification of
XX FcepsilonRI receptor and in the treatment and prophylaxis of allergic
XX diseases.
XX SQ Sequence 24 AA;

Query Match 43.0%; Score 58; DB 20; Length 24;
Best Local Similarity 55.6%; Pred. No. 0.021;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREV 19
   ||||| 1:11 ||| ::
DB 7 getyqcrvthphlpralm 24

RESULT 12
AAY50898
ID AAY50898 standard; peptide; 17 AA.
XX AC AAY50898;
XX DT 24-FEB-2000 (first entry)
XX DE Antibody 15A.2 swine IgE binding epitope 1.
XX KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
XX KW epitope; prophylaxis; treatment; mimotope.
XX OS Synthetic.
XX PN EP957111-A2.
XX PD 17-NOV-1999.
XX PF 09-APR-1999; 99EP-0107035.
XX PR 09-APR-1998; 98US-0058331.
XX PR 30-MAR-1999; 99US-0281760.
XX PA (IDEX-) IDEXX LAB INC.
XX PI Lawton R, Mermer B, Francoeur G;
XX PD WPI; 2000-040833/04.
```

XX Binding proteins used for treatment or prophylaxis of canine allergy -
PT Disclosure; Fig 7; 30pp; English.
PS This invention describes a novel binding protein which specifically
XX binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to IgE when the IgE is bound to mast cells. The peptide products of the
CC invention have anti-allergic activity. The antibodies bind to defined
CC epitopes on free or B-cell bound IgE molecules which have an important
CC role in allergic reaction. The specific binding proteins are used to
CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
XX invention.
XX SQ Sequence 17 AA;

Query Match 40.7%; Score 55; DB 21; Length 17;
Best Local Similarity 64.3%; Pred. No. 0.04;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 9 VSHPDLPREVRSI 22
Db |:||||:||||
4 vthpdkpiirsl 17

RESULT 13
AAY50893
ID AAY50893 standard; peptide; 17 AA.
XX
AC AAY50893;
XX
XX 24-FEB-2000 (first entry)
DE Antibody 15A.2 binding peptide 10 from PhdC7c phage display library.
XX Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
XX Synthetic.
XX EP957111-A2.
PN 17-NOV-1999.
XX
PD 09-APR-1999; 99EP-0107035.
XX
PF 09-APR-1998; 98US-0058331.
PR 30-MAR-1999; 99US-0281760.
XX
XX (IDEX-) IDEXX LAB INC.
XX
XX Lawton R, Mermer B, Francoeur G;
XX WPI; 2000-040833/04.
XX
XX Binding proteins used for treatment or prophylaxis of canine allergy -
PS Disclosure; Fig 6; 30pp; English.
XX
XX This invention describes a novel binding protein which specifically
CC binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to IgE when the IgE is bound to mast cells. The peptide products of the
CC invention have anti-allergic activity. The antibodies bind to defined
CC epitopes on free or B-cell bound IgE molecules which have an important
CC role in allergic reaction. The specific binding proteins are used to
CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
XX invention.
XX

SQ Sequence 17 AA;

Query Match 40.0%; Score 54; DB 21; Length 17;
Best Local Similarity 64.3%; Pred. No. 0.058;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 9 VSHPDLPREVRSI 22
Db |:||||:||||
4 vthpdkpiirsl 17

RESULT 14
AAY50894
ID AAY50894 standard; peptide; 17 AA.
XX
AC AAY50894;
XX
XX 24-FEB-2000 (first entry)
DE Antibody 15A.2 canine IgE binding epitope 1.
XX Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
XX Synthetic.
XX EP957111-A2.
PN 17-NOV-1999.
XX
PD 09-APR-1999; 99EP-0107035.
XX
PF 09-APR-1998; 98US-0058331.
PR 30-MAR-1999; 99US-0281760.
XX
XX (IDEX-) IDEXX LAB INC.
XX
XX Lawton R, Mermer B, Francoeur G;
XX WPI; 2000-040833/04.
XX
XX Binding proteins used for treatment or prophylaxis of canine allergy -
PS Disclosure; Fig 7; 30pp; English.
XX
XX This invention describes a novel binding protein which specifically
CC binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to IgE when the IgE is bound to mast cells. The peptide products of the
CC invention have anti-allergic activity. The antibodies bind to defined
CC epitopes on free or B-cell bound IgE molecules which have an important
CC role in allergic reaction. The specific binding proteins are used to
CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
XX invention.
XX SQ Sequence 17 AA;

Query Match 40.0%; Score 54; DB 21; Length 17;
Best Local Similarity 64.3%; Pred. No. 0.058;
Matches 9; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 9 VSHPDLPREVRSI 22
Db |:||||:||||
4 vthpdkpiirsl 17

RESULT 15
AAY50897
ID AAY50897 standard; peptide; 17 AA.
XX

AC AAY50897;
XX
XX
DT 24-FEB-2000 (first entry)
XX
DE Antibody 15A.2 feline IgE binding epitope 1.
XX
XX Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
XX
OS Synthetic.
XX
XX EP957111-A2.
FN
XX
PD 17-NOV-1999.
XX
XX 09-APR-1999; 99EP-0107035.
PF
XX
PR 09-APR-1998; 98US-0058331.
PR
XX 30-MAR-1999; 99US-0281760.
XX
XX (IDEX-) IDEXX LAB INC.
PA
XX
XX Lawton R, Mermer B, Francoeur G;
PI
XX
XX WPI; 2000-040333/04.
DR
XX
XX
PT Binding proteins used for treatment or prophylaxis of canine allergy -
PS Disclosure; Fig 7; 30pp; English.
XX
XX This invention describes a novel binding protein which specifically
CC binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to IgE when the IgE is bound to mast cells. The peptide products of the
CC invention have anti-allergic activity. The antibodies bind to defined
CC epitopes on free or B-cell bound IgE molecules which have an important
CC role in allergic reaction. The specific binding proteins are used to
CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAY50876-Y50300 represent peptide mimotopes used in the method of the
CC invention.
XX
XX
SQ Sequence 17 AA;

Query Match 39.3%; Score 53; DB 21; Length 17;
Best Local Similarity 71.4%; Pred No 0.084;
Matches 10; Conservative 2; Mismatches 2; Indels 0; Caps 0;

Qy 9 VSHPDLPFEVVRSI 22
I:||||| :|||
Db 4 vthpdlplvirsI 17

Search completed: March 4, 2002, 13:10:09
Job time: 706 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:10:08 : Search time 80.02 Seconds
(without alignments)
23.142 Million cell updates/sec

Title: US-09-701-623C-8
Perfect score: 140
Sequence: 1 CGYQSVIVDRDPFKPIVRSITLC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 220984

Minimum DB seq length: 0
Maximum DB seq length: 25

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 140 | 100.0 | 25 | 21 | AAV80001 |
| 2 | 106 | 75.7 | 25 | 21 | AAV80000 |
| 3 | 74 | 52.9 | 25 | 21 | AAV80077 |
| 4 | 66 | 47.1 | 25 | 21 | AAV79999 |
| 5 | 64 | 45.7 | 25 | 21 | AAV79998 |
| 6 | 64 | 45.7 | 25 | 21 | AAV91212 |
| 7 | 64 | 45.7 | 25 | 21 | AAV68602 |
| 8 | 54 | 38.6 | 15 | 21 | AAV50899 |
| 9 | 53 | 37.9 | 17 | 21 | AAV50898 |
| 10 | 45 | 32.1 | 17 | 11 | AAV02179 |
| 11 | 41 | 29.3 | 17 | 21 | AAV50893 |
| | | | | | Optimised IgE-CH3 |
| | | | | | Optimised IgE-CH3 |
| | | | | | Optimised IgE-CH3 |
| | | | | | Optimised IgE-CH3 |
| | | | | | Optimised IgE-CH3 |
| | | | | | Modified human IgE |
| | | | | | Peptide sequence o |
| | | | | | Antibody 15A.2 mur |
| | | | | | Antibody 15A.2 swi |
| | | | | | Antibody 15A.2 bin |

12 41 29.3 17 21 AAV50894
13 41 29.3 17 21 AAV50897
14 41 29.3 20 18 AAW24102
15 38.5 27.5 16 17 AAW09376
16 38.5 27.5 18 17 AAW09378
17 38 27.1 12 22 AAU15462
18 38 27.1 16 14 AAK30832
19 36 25.7 20 17 AAW05484
20 36 25.7 20 17 AAW05416
21 36 25.7 20 17 AAW16970
22 36 25.7 20 18 AAW25345
23 35.5 25.4 16 17 AAW09377
24 35 25.0 17 21 AAB20871
25 35 25.0 22 14 AAR41158
26 33.5 23.9 11 17 AAW17451
27 33 23.6 17 15 AAR51405
28 33 23.6 17 21 AAV50896
29 32 22.9 14 19 AAW59655
30 32 22.9 14 22 AAB88089
31 32 22.9 15 21 AAY98820
32 32 22.9 15 21 AAY98912
33 32 22.9 18 17 AAW05469
34 32 22.9 18 18 AAW37677
35 32 22.9 20 14 AAR41190
36 32 22.9 20 14 AAR41191
37 32 22.9 20 14 AAR41192
38 32 22.9 20 18 AAW10148
39 32 22.9 20 18 AAW10149
40 32 22.9 21 21 AAV58800
41 32 22.9 22 14 AAR41118
42 32 22.9 22 14 AAR41119
43 32 22.9 23 16 AAR70227
44 32 22.9 23 16 AAR70228
45 31 22.1 9 17 AAR89370

ALIGNMENTS

RESULT 1
AAV80001
ID AAV80001 standard; Peptide; 25 AA.

AC AAV80001;

DT 15-MAY-2000 (first entry)

XX Optimised IgE-CH3 domain antigen peptide for mouse IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Mus sp.

OS Synthetic.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy

XX Claim 1; Page 100; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils

CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

CC containing (I) are used for active immunisation against IgE-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell

CC target epitope, have increased immunogenicity and may include cyclic

CC constraints (disulfide bridge) to stabilise conformational features and

CC maximize cross-reactivity to the natural target. They induce safe

CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 25 AA;

Query Match 100.0%; Score 140; DB 21; Length 25;

Best Local Similarity 100.0%; Pred. No. 5.3e-15;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGYGYQSIIVDRPDPKPIVRSITLC 25

DB 1 CGYGYQSIIVDRPDPKPIVRSITLC 25

RESULT 2

AY80000

ID AAY80000 standard; Peptide; 25 AA.

AC AAY80000;

XX 15-MAY-2000 (first entry)

DE Optimised IgE-CH3 domain antigen peptide for rat IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

OS Rattus sp.

OS Synthetic.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy -

XX Claim 1; Page 99; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils

CC containing (I) are used for active immunisation against IgE-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell

CC target epitope, have increased immunogenicity and may include cyclic

CC constraints (disulfide bridge) to stabilise conformational features and

CC maximize cross-reactivity to the natural target. They induce safe

CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 25 AA;

Query Match 75.7%; Score 106; DB 21; Length 25;

Best Local Similarity 80.0%; Pred. No. 1.1e-09;

Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGYGYQSIIVDRPDPKPIVRSITLC 25

DB 1 CGYGYQSIIVDRPDPKPIVRSITLC 25

RESULT 3

AY80077

ID AAY80077 standard; Peptide; 25 AA.

XX AAY80077;

XX 15-MAY-2000 (first entry)

DE Optimised IgE-CH3 domain antigen peptide for horse IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

OS Equus caballus.

OS Synthetic.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy -

XX Claim 1; Page 146; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils

CC containing (I) are used for active immunisation against IgE-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;

Query Match 52.9%; Score 74; DB 21; Length 25;
 Best Local Similarity 52.0%; Pred. No. 0.0001;
 Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 CGYGQSIQVDRPDPKPIVRSITLC 25
 II I I I I I I I I I I I I I I I I
 Db 1 cgetyktvshpdlprevrslakc 25

RESULT 4
 AAY79999
 ID AAY79999 standard; Peptide; 25 AA.
 XX
 AC AAY79999;
 XX
 DT 15-MAY-2000 (first entry)
 XX
 DE Optimised IgE-CH3 domain antigen peptide for dog IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Canis sp.
 OS Synthetic.
 XX
 PN WO9967293-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13959.
 XX
 PR 20-JUN-1998; 98US-0100287.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY, Walfield AM;
 XX WPI; 2000-160578/14.
 XX
 DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy
 XX
 PS Claim 1; Page 99; 155pp; English.
 XX
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe

CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;

Query Match 47.1%; Score 66; DB 21; Length 25;
 Best Local Similarity 56.0%; Pred. No. 0.0018;
 Matches 14; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CGYGQSIQVDRPDPKPIVRSITLC 25
 II I I I I I I I I I I I I I I I I
 Db 1 cgetyysrvthphlpkdivrslakc 25

RESULT 5
 AAY79998
 ID AAY79998 standard; Peptide; 25 AA.
 XX
 AC AAY79998;
 XX
 DT 15-MAY-2000 (first entry)
 XX
 DE Optimised IgE-CH3 domain antigen peptide for human IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO9967293-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13959.
 XX
 PR 20-JUN-1998; 98US-0100287.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY, Walfield AM;
 XX WPI; 2000-160578/14.
 XX
 DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy
 XX
 PS Claim 1; Page 21; 155pp; English.
 XX
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;

Query Match 45.7%; Score 64; DB 21; Length 25;
 Best Local Similarity 48.0%; Pred. No. 0.0037;
 Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYQSVDRPDPFKPIVRSITLC 25
 II III I I I : : I I I
 Db 1 cgetyqsrvtphlpralmrsttkc 25

RESULT 6
 AAY91212
 ID AAY91212 standard; peptide; 25 AA.
 XX
 AC AAY91212;
 XX
 DT 22-MAY-2000 (first entry)
 DE Modified human IgE CH3 domain, SEQ ID NO:92.
 XX Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
 KW cholesteryl ester transport protein; anti-arteriosclerotic.
 XX Homo sapiens.
 OS Synthetic.
 XX WO9966957-A2.
 PN 29-DEC-1999.
 PD
 XX 21-JUN-1999; 99WO-US13975.
 PF
 XX 20-JUN-1998; 98US-0100412.
 PR
 XX (UNBI-) UNITED BIOMEDICAL INC.
 PA Wang CY;
 XX WPI: 2000-160564/14.
 DR
 XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus
 XX
 PS Example 6; Page 40; 129pp; English.
 XX
 CC The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesteryl ester transport
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration; for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences

CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVF Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVF Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
 CC protein epitope from *Yersinia* species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.
 XX
 SQ Sequence 25 AA;

Query Match 45.7%; Score 64; DB 21; Length 25;
 Best Local Similarity 48.0%; Pred. No. 0.0037;
 Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYQSVDRPDPFKPIVRSITLC 25
 II III I I I : : I I I
 Db 1 cgetyqsrvtphlpralmrsttkc 25

RESULT 7
 AAY68602
 ID AAY68602 standard; peptide; 25 AA.
 XX
 AC AAY68602;
 XX
 DT 05-MAY-2000 (first entry)
 DE Peptide sequence of the invention.
 XX
 KW Helper T cell epitope; peptide immunogen; LHRH;
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
 XX Unidentified.
 OS
 PN WO9966952-A1.
 XX 29-DEC-1999.
 PD
 XX 21-JUN-1999; 99WO-US13960.
 XX 20-JUN-1998; 98US-0100414.
 XX (UNBI-) UNITED BIOMEDICAL INC.
 XX Wang CY;
 XX WPI: 2000-160562/14.
 DR
 XX New peptide immunogen containing luteinising hormone-releasing hormone

PT antigen site and helper T cell epitope, for e.g. contraception and
 XX treatment of cancer -
 XX
 XX Disclosure; Page 92; 102pp; English.
 XX
 CC The specification describes peptide immunogens comprising a
 CC synthetic helper T cell (Th) epitope and a target antigen, luteinizing
 CC hormone-releasing hormone (LHRH). The peptide immunogens cause
 CC induction of a specific immune response to LHRH which is involved in
 CC regulation of spermatogenesis, ovulation, oestrus, sexual development
 CC and secretion of sex hormones. Provision of a promiscuous T helper
 CC epitope (which is functional in genetically diverse subjects) provides
 CC optimum immunogenicity to the B cell epitopes of the target antigen and
 CC thus high antibody titres against the target antigen. The peptide
 CC immunogens of the invention are used to vaccinate against mammalian LHRH,
 CC for use as (reversible) contraceptive; control of hormone-dependent
 CC tumours (cancer of prostate or breast; also endometriosis); to prevent
 CC boar taint (and improve meat quality) and for immunocastration. The
 CC present sequence appears in the specification.
 XX
 XX Sequence 25 AA;
 SQ

Query Match 45.7%; Score 64; DB 21; Length 25;
 Best Local Similarity 48.0%; Pred. No. 0.0037;
 Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYQIVDRDPFKPIVRSITLC 25
 || ||| | | | : || | |
 Db 1 cgetygsrvthplpralmrsttkc 25

RESULT 8
 AAY50899
 ID AAY50899 standard; peptide; 15 AA.
 XX
 AC AAY50899;
 XX
 XX 24-FEB-2000 (first entry)
 XX
 DE Antibody 15A.2 murine IgE binding epitope 1.
 XX
 KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 KW epitope; prophylaxis; treatment; mimotope.
 XX
 OS Synthetic.
 XX
 PN EP957111-A2.
 XX
 PD 17-NOV-1999.
 XX
 PF 09-APR-1999; 99EP-0107035.
 XX
 PR 09-APR-1998; 98US-0058331.
 PR 30-MAR-1999; 99US-0281760.
 XX
 PA (IDEX-) IDEXX LAB INC.
 XX
 PI Lawton R, Mermer B, Francoeur G;
 XX
 XX WPI; 2000-040833/04.
 XX
 PD Binding proteins used for treatment or prophylaxis of canine allergy -
 XX Disclosure; Fig 7; 30pp; English.
 XX
 PF This invention describes a novel binding protein which specifically
 XX binds to native canine free or B-cell bound IgE, and which doesn't bind
 XX to IgE when the IgE is bound to mast cells. The peptide products of the
 XX invention have anti-allergic activity. The antibodies bind to defined
 XX epitopes on free or B-cell bound IgE molecules which have an important
 XX role in allergic reaction. The specific binding proteins are used to
 XX produce a pharmaceutical composition, preferably with a diluent, which
 XX can be used for prophylaxis or treatment of canine allergy.
 XX
 XX AAY50876-Y50900 represent peptide mimotopes used in the method of the
 XX invention.
 XX
 XX Sequence 17 AA;
 SQ

Query Match 37.9%; Score 53; DB 21; Length 17;
 Best Local Similarity 71.4%; Pred. No. 0.12;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 VDRPDFPKPIVRSI 22
 | || ||| |||||
 Db 4 vthpdlpkpilsrl 17

CC can be used for prophylaxis or treatment of canine allergy.
 CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.
 XX
 XX Sequence 15 AA;
 SQ

Query Match 38.6%; Score 54; DB 21; Length 15;
 Best Local Similarity 76.5%; Pred. No. 0.075;
 Matches 13; Conservative 0; *Mismatches 2; Indels 2; Gaps 1;

QY 6 QSIIVDRDPFKPIVRSI 22
 | ||| ||| |||||
 Db 1 qcivdhpdf--pivrsi 15

RESULT 9
 AAY50898
 ID AAY50898 standard; peptide; 17 AA.
 XX
 AC AAY50898;
 XX
 XX 24-FEB-2000 (first entry)
 XX
 DE Antibody 15A.2 swine IgE binding epitope 1.
 XX
 KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 KW epitope; prophylaxis; treatment; mimotope.
 XX
 OS Synthetic.
 XX
 PN EP957111-A2.
 XX
 PD 17-NOV-1999.
 XX
 PF 09-APR-1999; 99EP-0107035.
 XX
 PR 09-APR-1998; 98US-0058331.
 PR 30-MAR-1999; 99US-0281760.
 XX
 PA (IDEX-) IDEXX LAB INC.
 XX
 PI Lawton R, Mermer B, Francoeur G;
 XX
 XX WPI; 2000-040833/04.
 XX
 PD Binding proteins used for treatment or prophylaxis of canine allergy -
 XX Disclosure; Fig 7; 30pp; English.
 XX
 PF This invention describes a novel binding protein which specifically
 XX binds to native canine free or B-cell bound IgE, and which doesn't bind
 XX to IgE when the IgE is bound to mast cells. The peptide products of the
 XX invention have anti-allergic activity. The antibodies bind to defined
 XX epitopes on free or B-cell bound IgE molecules which have an important
 XX role in allergic reaction. The specific binding proteins are used to
 XX produce a pharmaceutical composition, preferably with a diluent, which
 XX can be used for prophylaxis or treatment of canine allergy.
 XX
 XX AAY50876-Y50900 represent peptide mimotopes used in the method of the
 XX invention.
 XX
 XX Sequence 17 AA;
 SQ

Query Match 37.9%; Score 53; DB 21; Length 17;
 Best Local Similarity 71.4%; Pred. No. 0.12;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 9 VDRPDFPKPIVRSI 22
 | || ||| |||||
 Db 4 vthpdlpkpilsrl 17

RESULT 10
 AAR02179
 ID AAR02179 standard; protein; 17 AA.
 XX
 AC AAR02179;
 XX
 DT 02-AUG-1990 (first entry)
 XX
 DE Peptide which binds to Immunoglobulin E (IgE) receptors inhibiting IgE
 DE binding.
 XX
 KW Immunoglobulin E (IgE); anti-allergic drug; IgE binding inhibitor.
 XX
 FH Key Location/Qualifiers
 FT MISC-difference 1 /label-OTHER
 FT /note="H-Pro"
 FT MISC-difference 17
 FT /label-OTHER
 FT /note="Ala-OH"
 XX
 PN JP01299298-A.
 XX
 PD 04-DEC-1989.
 XX
 PF 27-MAY-1988; 88JP-0129721.
 XX
 PR 27-MAY-1988; 88JP-0129721.
 XX
 PA (KYOW) KYOWA HAKKO KOGYO KK.
 XX
 DR WPI; 1990-019330/03.
 XX
 PT Peptides which bind to IgE receptors inhibiting IgE binding -
 PT used as anti-allergic drug having no side effects
 XX
 PS Claim 2; Page 843; 8pp; Japanese.
 XX
 CC It is the new sequence No 11 and has an anti-allergic effect. It binds to
 CC IgE receptors on cell surfaces and inhibits IgE binding to receptors. It
 CC can be used as an anti-allergic drug, having no side effects.
 XX
 SQ Sequence 17 AA;
 Query Match 32.1%; Score 45; DB 11; Length 17;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 15 PKPIVRSIT 23
 DB 1 pkpivruit 9
 RESULT 11
 AAY50893
 ID AAY50893 standard; peptide; 17 AA.
 XX
 AC AAY50893;
 XX
 DT 24-FEB-2000 (first entry)
 XX
 DE Antibody 15A.2 binding peptide 10 from Phbdc7c phage display library.
 XX
 KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 KW epitope; prophylaxis; treatment; mimotope.
 XX
 OS Synthetic.
 XX
 PN EP957111-A2.
 XX
 PD 17-NOV-1999.
 XX
 QY 15 PKPIVRSIT 23
 DB 1 pkpivruit 9

PF 09-APR-1999; 99EP-0107035.
 XX
 PR 09-APR-1998; 98US-0058331.
 PR 30-MAR-1999; 99US-0281760.
 XX
 PA (IDEX-) IDEXX LAB INC.
 XX
 PI Lawton R, Mermer B, Francoeur G;
 XX
 DR WPI; 2000-040833/04.
 XX
 PT Binding proteins used for treatment or prophylaxis of canine allergy -
 XX
 PS Disclosure; Fig 6; 30pp; English.
 XX
 CC This invention describes a novel binding protein which specifically
 CC binds to native canine free or B-cell bound IgE, and which doesn't bind
 CC to IgE when the IgE is bound to mast cells. The peptide products of the
 CC invention have anti-allergic activity. The antibodies bind to defined
 CC epitopes on free or B-cell bound IgE molecules which have an important
 CC role in allergic reaction. The specific binding proteins are used to
 CC produce a pharmaceutical composition, preferably with a diluent, which
 CC can be used for prophylaxis or treatment of canine allergy.
 CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.
 XX
 SQ Sequence 17 AA;
 Query Match 29.3%; Score 41; DB 21; Length 17;
 Best Local Similarity 64.3%; Pred. No. 9.1;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 9 VDRPDFPKPIVRSI 22
 DB 4 vtphlpkdivrsi 17
 RESULT 12
 AAY50894
 ID AAY50894 standard; peptide; 17 AA.
 XX
 AC AAY50894;
 XX
 DT 24-FEB-2000 (first entry)
 XX
 DE Antibody 15A.2 canine IgE binding epitope 1.
 XX
 KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 KW epitope; prophylaxis; treatment; mimotope.
 XX
 OS Synthetic.
 XX
 PN EP957111-A2.
 XX
 PD 17-NOV-1999.
 XX
 PF 09-APR-1999; 99EP-0107035.
 XX
 PR 09-APR-1998; 98US-0058331.
 PR 30-MAR-1999; 99US-0281760.
 XX
 PA (IDEX-) IDEXX LAB INC.
 XX
 PI Lawton R, Mermer B, Francoeur G;
 XX
 DR WPI; 2000-040833/04.
 XX
 PT Binding proteins used for treatment or prophylaxis of canine allergy -
 XX
 PS Disclosure; Fig 7; 30pp; English.
 XX
 CC This invention describes a novel binding protein which specifically

CC binds to native canine free or B-cell bound IgE, and which doesn't bind
 CC to IgE when the IgE is bound to mast cells. The peptide products of the
 CC invention have anti-allergic activity. The antibodies bind to defined
 CC epitopes on free or B-cell bound IgE molecules which have an important
 CC role in allergic reaction. The specific binding proteins are used to
 CC produce a pharmaceutical composition, preferably with a diluent, which
 CC can be used for prophylaxis or treatment of canine allergy.
 CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.
 XX
 XX Sequence 17 AA;

Query Match 29.3%; Score 41; DB 21; Length 17;
 Best Local Similarity 64.3%; Pred. No. 9.1;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 VDRPDPFKPIVRSI 22
 | | | | |
 Db 4 vthphlpkdivrsi 17

RESULT 13
 AAY50897
 ID AAY50897 standard; peptide; 17 AA.
 XX
 XX AC AAY50897;
 XX
 XX 24-FEB-2000 (first entry)
 XX
 XX DE Antibody 15A.2 feline IgE binding epitope 1.
 XX
 XX Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 KW epitope; prophylaxis; treatment; mimotope.
 KW
 OS Synthetic.
 OS
 XX EP957111-A2.
 PN
 XX 17-NOV-1999.
 PD
 XX
 XX 09-APR-1999; 99EP-0107035.
 PF
 XX 09-APR-1998; 98US-0058331.
 PR
 XX 30-MAR-1999; 99US-0281760.
 XX
 XX (IDEX-) IDEXX LAB INC.
 PA
 XX Lawton R, Mermer B, Francoeur G;
 PI
 XX WPI; 2000-040833/04.
 DR
 XX Binding proteins used for treatment or prophylaxis of canine allergy -
 XX Disclosure; Fig 7; 30pp; English.

CC This invention describes a novel binding protein which specifically
 CC binds to native canine free or B-cell bound IgE, and which doesn't bind
 CC to IgE when the IgE is bound to mast cells. The peptide products of the
 CC invention have anti-allergic activity. The antibodies bind to defined
 CC epitopes on free or B-cell bound IgE molecules which have an important
 CC role in allergic reaction. The specific binding proteins are used to
 CC produce a pharmaceutical composition, preferably with a diluent, which
 CC can be used for prophylaxis or treatment of canine allergy.
 CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.
 XX
 XX Sequence 17 AA;

Query Match 29.3%; Score 41; DB 21; Length 17;
 Best Local Similarity 58.8%; Pred. No. 9.1;
 Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 6 QSTVDRPDPFKPIVRSI 22
 | | | | |
 Db 1 qckvthpdpkplvirs 17

RESULT 14
 AAW24102
 ID AAW24102 standard; peptide; 20 AA.
 XX
 XX AC AAW24102;
 XX
 XX 21-NOV-1997 (first entry)
 DT
 XX Canine immunoglobulin E peptide 5.
 DE
 XX Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.
 KW
 XX Canis familiaris.
 OS
 XX JP09169795-A.
 PN
 XX 30-JUN-1997.
 PD
 XX 22-DEC-1995; 95JP-0334381.
 PF
 XX 22-DEC-1995; 95JP-0334381.
 PR
 XX (HITB) HITACHI CHEM CO LTD.
 PA
 XX WPI; 1997-389423/36.
 DR
 XX N-PSDB; AAT85651.
 DR
 XX Canine immunoglobulin E peptide fragment and related DNA - useful
 PT for the preparation of anti-canine immunoglobulin E antibody
 PT
 XX Claim 2; Page 9; 12pp; Japanese.

CC AAW24098-106 are peptide fragments containing at least 5 continuous
 CC amino acids of the partial canine immunoglobulin E (IgE) protein shown
 CC in AAW24097. The peptides are used for the preparation of anti-canine
 CC IgE antibody. The anti-canine IgE antibody can be used for the diagnosis
 CC of canine allergies.
 XX
 XX Sequence 20 AA;

Query Match 29.3%; Score 41; DB 18; Length 20;
 Best Local Similarity 64.3%; Pred. No. 11;
 Matches 9; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 VDRPDPFKPIVRSI 22
 | | | | |
 Db 4 vthphlpkdivrsi 17

RESULT 15
 AAW09376
 ID AAW09376 standard; peptide; 16 AA.
 XX
 XX AC AAW09376;
 AC
 XX 17-JUN-1997 (first entry)
 DT
 XX Proline-rich antibacterial peptide from Podisus maculiventris.
 DE
 XX Podisus maculiventris; proline-rich; pathogenic; bacteria; bug;
 KW Gram negative; plant treatment.
 KW
 XX Podisus maculiventris.
 OS
 XX FR2732345-A1.
 PN
 XX

PD 04-OCT-1996.
 XX
 PF 03-APR-1995; 95PR-0004130.
 XX
 PR 03-APR-1995; 95PR-0004130.
 XX
 PA (RHON) RHONE POULENC AGROCHIMIE.
 XX
 PI Bulet P, Hetru C, Hoffman J, Tchernych S;
 XX
 DR WPI; 1996-457711/46.
 XX
 PT Proline-rich antibacterial peptide(s) from Podisus maculiventris -
 PT esp. useful to protect plants against pathogenic bacteria, also for
 PT antibacterial therapy of animals
 XX
 PS Claim 2; Page 14; 16pp; English.
 XX
 CC AAW09375-W09378 are specific examples of antibacterial peptides from the
 CC bug Podisus maculiventris. The proline-rich peptides have antibacterial
 CC activity against gram negative bacteria. They are useful for protecting
 CC plants against pathogenic bacteria, but could also be used for
 CC antibacterial therapy of animals, including humans. A generic peptide
 CC having the sequence VDKPDRPRPX is also given in the specification,
 CC where x is a peptide comprising at least one PRP motif (see AAW17451).
 XX
 SQ Sequence 16 AA;
 XX
 Query Match 27.5%; Score 38.5; DB 17; Length 16;
 Best Local Similarity 43.8%; Pred. No. 21;
 Matches 7; Conservative 4; Mismatches 0; Indels 5; Gaps 1;
 QY 9 VDRPDF-----PKPIV 19
 Db |||||
 1 vdkpdyrprgprpi 16
 Search completed: March 4, 2002, 13:10:09
 Job time: 706 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:10:08 ; Search time 80.02 seconds
(without alignments)
23.142 Million cell updates/sec

Title: US-09-701-623C-7

Perfect score: 145

Sequence: 1 CGEGYQRVDHPHPRIVRSYK 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 220984

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|----------------------|
| 1 | 145 | 100.0 | 25 | 21 AAY80000 | Optimised IgE-CH3 |
| 2 | 106 | 73.1 | 25 | 21 AAY80001 | Optimised IgE-CH3 |
| 3 | 105 | 72.4 | 25 | 21 AAY79999 | Optimised IgE-CH3 |
| 4 | 103 | 71.0 | 25 | 21 AAY79998 | Optimised IgE-CH3 |
| 5 | 103 | 71.0 | 25 | 21 AAY91212 | Modified human IgE |
| 6 | 103 | 71.0 | 25 | 21 AAY68602 | Peptide sequence o |
| 7 | 89 | 61.4 | 25 | 21 AAY80077 | Optimised IgE-CH3 |
| 8 | 68 | 46.9 | 20 | 18 AAW24102 | Canine immunoglobuli |
| 9 | 63 | 43.4 | 17 | 21 AAY50893 | Antibody 15A.2 bin |
| 10 | 63 | 43.4 | 17 | 21 AAY50894 | Antibody 15A.2 can |
| 11 | 62 | 42.8 | 22 | 20 AAY42585 | IgE peptide antago |

| | | | | | | |
|----|------|------|----|----|----------|--------------------|
| 12 | 52 | 42.8 | 24 | 20 | AAY42586 | IgE peptide antago |
| 13 | 58 | 40.0 | 17 | 21 | AAY50896 | Antibody 15A.2 gre |
| 14 | 55 | 37.9 | 17 | 21 | AAY50895 | Antibody 15A.2 hum |
| 15 | 54 | 37.2 | 17 | 21 | AAY50898 | Antibody 15A.2 swi |
| 16 | 51 | 35.2 | 25 | 21 | AAY80048 | IgE derived target |
| 17 | 50 | 34.5 | 17 | 11 | AAR02179 | Peptide which bind |
| 18 | 48 | 33.1 | 15 | 21 | AAY50899 | Antibody 15A.2 mur |
| 19 | 47 | 32.4 | 17 | 21 | AAY50897 | Antibody 15A.2 fel |
| 20 | 44 | 30.3 | 19 | 22 | AAB26503 | Human IgE C epsilo |
| 21 | 44 | 30.3 | 19 | 22 | AAB51033 | IgE peptide #11. |
| 22 | 43 | 29.7 | 13 | 21 | AAB26518 | Human IgE C epsilo |
| 23 | 41 | 28.3 | 22 | 16 | AAR72686 | Peptide #3 having |
| 24 | 39 | 26.9 | 22 | 16 | AAR72687 | Peptide #4 having |
| 25 | 37.5 | 25.9 | 18 | 22 | AAB50754 | Human CAMP-specifi |
| 26 | 37 | 25.5 | 15 | 18 | AAM32537 | Anti-amyloid pepti |
| 27 | 36 | 24.8 | 25 | 22 | AAM36513 | Peptide #10550 enc |
| 28 | 35 | 24.1 | 9 | 21 | AAY69599 | Monoclonal antibod |
| 29 | 35 | 24.1 | 9 | 21 | AAY69600 | Monoclonal antibod |
| 30 | 35 | 24.1 | 9 | 21 | AAY50891 | Antibody 15A.2 bin |
| 31 | 35 | 24.1 | 9 | 21 | AAY50892 | Antibody 15A.2 bin |
| 32 | 35 | 24.1 | 14 | 21 | AAB26519 | Human IgE C epsilo |
| 33 | 35 | 24.1 | 14 | 21 | AAY69579 | Immunogenic peptid |
| 34 | 34.5 | 23.8 | 12 | 21 | AAB26517 | Human IgE C epsilo |
| 35 | 34.5 | 23.8 | 16 | 21 | AAB24095 | Human apoptosis re |
| 36 | 34 | 23.4 | 12 | 21 | AAB26516 | Human IgE C epsilo |
| 37 | 34 | 23.4 | 12 | 22 | AAB51038 | IgE peptide #16. |
| 38 | 34 | 23.4 | 13 | 20 | AAY48825 | Membrane dipeptida |
| 39 | 34 | 23.4 | 15 | 21 | AAB69220 | HIV-1 non-subtype |
| 40 | 34 | 23.4 | 18 | 18 | AAM38974 | Peptide resembling |
| 41 | 34 | 23.4 | 20 | 11 | AAR04962 | Papilloma virus ty |
| 42 | 33.5 | 23.1 | 15 | 21 | AAY98912 | HLA class II bindi |
| 43 | 33 | 22.8 | 17 | 18 | AAM39002 | Peptide resembling |
| 44 | 33 | 22.8 | 17 | 18 | AAM39001 | Peptide resembling |
| 45 | 33 | 22.8 | 19 | 9 | AAP81947 | 19 amino acid pept |

ALIGNMENTS

RESULT 1
AAY80000
ID AAY80000 standard; Peptide; 25 AA.
XX
AC AAY80000;
XX
DT 15-MAY-2000 (first entry)
XX
DE Optimised IgE-CH3 domain antigen peptide for rat IgE.
XX
KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
OS Rattus sp.
OS Synthetic.
XX
PN WO967293-AL.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13959.
XX
PR 20-JUN-1998; 98US-0100287.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY, Walfield AM;
XX
WP; 2000-160578/14.
XX
PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
for immunization against allergy

CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;

Query Match 72.4%; Score 105; DB 21; Length 25;
 Best Local Similarity 76.0%; Pred. No. 4.8e-09;
 Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
 ||| | ||| ||| ||| ||| ||
 Db 1 cgetysrvthphlpxdivrslakc 25

RESULT 4

AAV79998
 ID AAY79998 standard; Peptide; 25 AA.

XX
 AC AAY79998;

DT 15-MAY-2000 (first entry)

DE Optimised IgE-CH3 domain antigen peptide for human IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX

OS Homo sapiens.
 OS Synthetic.

XX WO9967293-A1.

PN 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -
 XX

XX Claim 1; Page 21; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe

CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;

Query Match 71.0%; Score 103; DB 21; Length 25;
 Best Local Similarity 68.0%; Pred. No. 9.6e-09;
 Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
 ||| ||||| ||| | : || |||
 Db 1 cgetysrvthphlpraimrsttkc 25

RESULT 5

AAV91212

ID AAY91212 standard; peptide; 25 AA.

XX
 AC AAY91212;

XX 22-MAY-2000 (first entry)

DE Modified human IgE CH3 domain, SEQ ID NO:92.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
 KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX Homo sapiens.
 OS Synthetic.

XX WO9966957-A2.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13975.

XX 20-JUN-1998; 98US-0100412.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI; 2000-160564/14.

XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus -
 XX

XX Example 6; Page 40; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesteryl ester transport
 CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infectious (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration); for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target

CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVP) protein and sequences AAY91122-Y91142.
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVP Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVP Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CEMP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CEMP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitopes which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
 CC protein epitope from *Yersinia* species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

XX
 SQ Sequence 25 AA;

Query Match 71.0%; Score 103; DB 21; Length 25;
 Best Local Similarity 68.0%; Pred. No. 9.6e-09;
 Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPPIVRSITKNC 25
 DB 1 cgetyqsrvtphlpralmrstktc 25

RESULT 6
 AAY68602
 ID AAY68602 standard; peptide; 25 AA.

XX
 AC AAY68602;

DT 05-MAY-2000 (first entry)

XX Peptide sequence of the invention.

XX Helper T cell epitope; peptide immunogen; LHRH;
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.

XX Unidentified.

OS

XX WO9966952-A1.

PN 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13960.

XX 20-JUN-1998; 98US-0100414.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;
 PI
 XX WPI; 2000-160562/14.

XX New peptide immunogen containing luteinising hormone-releasing hormone
 PT antigen site and helper T cell epitope, for e.g. contraception and
 PT treatment of cancer

XX Disclosure; Page 92; 102pp; English.

XX The specification describes peptide immunogens comprising a
 CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
 CC hormone-releasing hormone (LHRH). The peptide immunogens cause
 CC induction of a specific immune response to LHRH which is involved in
 CC regulation of spermatogenesis, ovulation, oestrus, sexual development
 CC and secretion of sex hormones. Provision of a promiscuous T helper
 CC epitope (which is functional in genetically diverse subjects) provides
 CC optimum immunogenicity to the B cell epitopes of the target antigen and
 CC thus high antibody titres against the target antigen. The peptide
 CC immunogens of the invention are used to vaccinate against mammalian LHRH,
 CC for use as (reversible) contraceptive; control of hormone-dependent
 CC tumours (cancer of prostate or breast, also endometriosis); to prevent
 CC boar taint (and improve meat quality) and for immunocastration. The
 CC present sequence appears in the specification.

XX Sequence 25 AA;

Query Match 71.0%; Score 103; DB 21; Length 25;
 Best Local Similarity 68.0%; Pred. No. 9.6e-09;
 Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPPIVRSITKNC 25
 DB 1 cgetyqsrvtphlpralmrstktc 25

RESULT 7
 AAY80077
 ID AAY80077 standard; Peptide; 25 AA.

XX AAY80077;

DT 15-MAY-2000 (first entry)

XX Optimised IgE-CH3 domain antigen peptide for horse IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Equus caballus.

OS Synthetic.

XX WO9967293-A1.

PN 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy

```

Query Match      46.9%; Score 68; DB 18; Length 20;
Best Local Similarity 76.5%; Pred. NO. 0.0014;
Matches 13; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      8 RVDHPHFPPKPIVRISIK 24
         || ||| || ||||| |
Ddb     3 rvthphlpkdivrsiak 19

RESULT      9
ID AAAY50893
XX AAAY50893 standard; peptide; 17 AA.
XX
XX AAAY50893;
XX
XX
DT 24-FEB-2000 (first entry)
XX
XX Antibody 15A.2 binding peptide 10 from PhDc7c phage display library.
DE
XX
XX Canine; allergy; antibody 15A.2; IGE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
KW

```

OS
XX
synthetic.
PN
EP957111-A2.
XX
XX
17-NOV-1999.
PD
XX
XX
09-APR-1999; 99EP-0107035.
XX
XX

30 MAR 1993, 9303 0261760.
(INDEX-) IDEXX LAB INC.
Lawton R, Meirner B, Francoeur G;
WPI; 2000-040833/04.
Binding proteins used for treatment or prophylaxis of canine allergy -
Disclosure; Fig 6; 30pp; English.

| | |
|-----------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CC | This invention describes a novel binding protein which specifically binds to native canine free or B-cell bound IgE, and which doesn't bind to IgE when the IgE is bound to mast cells. The peptide products of the invention have anti-allergic activity. The antibodies bind to defined epitopes on free or B-cell bound IgE molecules which have an important role in allergic reaction. The specific binding proteins are used to produce a pharmaceutical composition, preferably with a diluent, which can be used for prophylaxis or treatment of canine allergy. |
| CC | AAY50876-Y50900 represent peptide mimotopes used in the method of the invention. |
| XX | |
| SQ | Sequence 17 AA; |
| Query Match | 43.4%; Score 63; DB 21; Length 17; |
| Best Local Similarity | 80.0%; Pred. No. 0.0064; |
| Matches | 12; Conservative 0; Mismatches 3; Indels 0; Gaps 0; |
| OY | 8 RVDHPHFPKPIVRSI 22 |
| Db | 3 rvthphlpkdivrsi 17 |
| RESULT | 10 |
| AAY50894 | |
| ID | AAY50894 standard; peptide; 17 AA. |

XX
AC AAY50894:

XX
DT 24-FEB-2000 (first entry)

PT diseases -

XX PS Disclosure; Column 9; 37pp; English.

XX CC The invention provides immunoglobulin E (IgE) antagonists comprising one or more of the FcεpsilonRI receptor-binding determinant sites of human IgE. The antagonists include IgE variants comprising an immunoglobulin template and binding determinant sequences (bds) Cbds, Efbds and the sequence shown in AAY42581. The Cbds (CD loop binding determinant sequence) are selected from the sequences shown in AAY42567-Y42577 and the Efbds (EF loop binding determinant sequence) are selected from sequences shown in AAY42578-Y42580. The variants are useful in raising and screening anti-IgE antibodies, in the isolation and purification of FcεpsilonRI receptor and in the treatment and prophylaxis of allergic diseases.

XX SQ Sequence 24 AA;

Query Match 42.8%; Score 62; DB 20; Length 24;

XX Best Local Similarity 55.6%; Pred. No. 0.013; Mismatches 3; Indels 0; Gaps 0;

XX Mates 10; Conservative 3;

QY 2 GEGYQSRVDHPHPKPIV 19

DB 7 getyqcrvthphlpralm 24

RESULT 13

AAY50896

ID AAY50896 standard; peptide; 17 AA.

XX AC AAY50896;

XX DT 24-FEB-2000 (first entry)

XX DE Antibody 15A.2 green monkey IgE binding epitope 1.

XX KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy; epitope; prophylaxis; treatment; mimotope.

XX OS Synthetic.

XX PN EP957111-A2.

XX PD 17-NOV-1999.

XX PF 09-APR-1999; 99EP-0107035.

XX PR 09-APR-1998; 98US-0058331.

XX PR 30-MAR-1999; 99US-0281760.

XX PA (IDEX-) IDEXX LAB INC.

XX PI Lawton R, Mermer B, Francoeur G;

XX DR WPI; 2000-040833/04.

XX Binding proteins used for treatment or prophylaxis of canine allergy -

XX PS Disclosure; Fig 7; 30pp; English.

XX CC This invention describes a novel binding protein which specifically binds to native canine free or B-cell bound IgE, and which doesn't bind to IgE when the IgE is bound to mast cells. The peptide products of the invention have anti-allergic activity. The antibodies bind to defined epitopes on free or B-cell bound IgE molecules which have an important role in allergic reaction. The specific binding proteins are used to produce a pharmaceutical composition, preferably with a diluent, which can be used for prophylaxis or treatment of canine allergy.

XX CC AAY50876-Y50900 represent peptide mimotopes used in the method of the invention.

XX

SQ Sequence 17 AA;

Query Match 40.0%; Score 58; DB 21; Length 17;

XX Best Local Similarity 62.5%; Pred. No. 0.036; Mismatches 2; Indels 0; Gaps 0;

XX Mates 10; Conservative 2;

QY 6 QSRVDHPHPKPIVRS 21

DB 1 qcrvthphlpralvrs 16

RESULT 14

AAY50895

ID AAY50895 standard; peptide; 17 AA.

XX AC AAY50895;

XX DT 24-FEB-2000 (first entry)

XX DE Antibody 15A.2 human IgE binding epitope 1.

XX KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy; epitope; prophylaxis; treatment; mimotope.

XX OS Synthetic.

XX PN EP957111-A2.

XX PD 17-NOV-1999.

XX PF 09-APR-1999; 99EP-0107035.

XX PR 09-APR-1998; 98US-0058331.

XX PR 30-MAR-1999; 99US-0281760.

XX PA (IDEX-) IDEXX LAB INC.

XX PI Lawton R, Mermer B, Francoeur G;

XX DR WPI; 2000-040833/04.

XX Binding proteins used for treatment or prophylaxis of canine allergy -

XX PS Disclosure; Fig 7; 30pp; English.

XX CC This invention describes a novel binding protein which specifically binds to native canine free or B-cell bound IgE, and which doesn't bind to IgE when the IgE is bound to mast cells. The peptide products of the invention have anti-allergic activity. The antibodies bind to defined epitopes on free or B-cell bound IgE molecules which have an important role in allergic reaction. The specific binding proteins are used to produce a pharmaceutical composition, preferably with a diluent, which can be used for prophylaxis or treatment of canine allergy.

XX CC AAY50876-Y50900 represent peptide mimotopes used in the method of the invention.

XX SQ Sequence 17 AA;

Query Match 37.9%; Score 55; DB 21; Length 17;

XX Best Local Similarity 56.2%; Pred. No. 0.1; Mismatches 3; Indels 0; Gaps 0;

XX Mates 9; Conservative 3;

QY 6 QSRVDHPHPKPIVRS 21

DB 1 qcrvthphlpralmrs 16

RESULT 15

AAY50898

ID AAY50898 standard; peptide; 17 AA.

XX

AC AAY50898;
 XX
 XX 24-FEB-2000 (first entry)
 XX
 DE Antibody 15A.2 swine IgE binding epitope 1.
 XX
 XX Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
 KW epitope; prophylaxis; treatment; mimotope.
 XX
 XX Synthetic.
 OS
 XX EP957111-A2.
 PN
 XX 17-NOV-1999.
 PD
 XX
 XX 09-APR-1999; 99EP-0107035.
 PF
 XX 09-APR-1998; 98US-0058331.
 PR 30-MAR-1999; 99US-0281760.
 XX
 XX (IDEX-) IDEX LAB INC.
 PA
 XX Lawton R, Marmer B, Francoeur G;
 PI
 XX WPI; 2000-040833/04.
 DR
 XX
 XX Binding proteins used for treatment or prophylaxis of canine allergy -
 PT
 XX Disclosure; Fig 7; 30pp; English.
 PS
 XX This invention describes a novel binding protein which specifically
 CC binds to native canine free or B-cell bound IgE, and which doesn't bind
 CC to IgE when the IgE is bound to mast cells. The peptide products of the
 CC invention have anti-allergic activity. The antibodies bind to defined
 CC epitopes on free or B-cell bound IgE molecules which have an important
 CC role in allergic reaction. The specific binding proteins are used to
 CC produce a pharmaceutical composition, preferably with a diluent, which
 CC can be used for prophylaxis or treatment of canine allergy.
 CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
 CC invention.
 CC
 XX Sequence 17 AA;
 SQ

Query Match 37.2%; Score 54; DB 21; Length 17;
 Best Local Similarity 71.4%; Pred. NO. 0.14;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 9 VDRPHFPKPIVRSI 22
 | | | | | | | | | |
 Db 4 vthpdlpkpilsr 17

Search completed: March 4, 2002, 13:10:08
 Job time: 705 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:10:08 ; Search time 80.02 Seconds
(without alignments)
23.142 Million cell updates/sec

Title: US-09-701-623c-6

Perfect score: 141

Sequence: 1 GGTYYRVTTHPLPKDIVRSIAKC 25

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 220984

Minimum DB seq length: 0

Maximum DB seq length: 25

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_1101.*

```
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 141 | 100.0 | 25 | 21 AAY79999 | Optimised IgE-CH3 |
| 2 | 108 | 76.6 | 25 | 21 AAY79998 | Optimised IgE-CH3 |
| 3 | 108 | 76.6 | 25 | 21 AAY91212 | Modified human IgE |
| 4 | 108 | 76.6 | 25 | 21 AAY68602 | Peptide sequence o |
| 5 | 105 | 74.5 | 25 | 21 AAY80000 | Optimised IgE-CH3 |
| 6 | 105 | 74.5 | 25 | 21 AAY80077 | Optimised IgE-CH3 |
| 7 | 95 | 67.4 | 20 | 18 AAW24102 | Canine immunoglobi |
| 8 | 86 | 61.0 | 17 | 21 AAY50893 | Antibody 15A.2 bin |
| 9 | 86 | 61.0 | 17 | 21 AAY50894 | Antibody 15A.2 can |
| 10 | 72 | 51.1 | 22 | 20 AAY42585 | IgE peptide antago |
| 11 | 72 | 51.1 | 24 | 20 AAY42586 | IgE peptide antago |

```
12 66 46.8 25 21 AAY80001 Optimised IgE-CH3
13 63 44.7 17 21 AAY50896 Antibody 15A.2 gre
14 62 44.0 17 21 AAY50898 Antibody 15A.2 swi
15 60 42.6 17 21 AAY50895 Antibody 15A.2 hum
16 58 41.1 25 21 AAY80048 IgE derived target
17 52 36.9 17 21 AAY50897 Antibody 15A.2 fel
18 51 36.2 13 21 AAB26518 Human IgE C epsilo
19 43.5 30.9 12 21 AAB26517 Human IgE C epsilo
20 42 29.8 19 21 AAB26503 Human IgE C epsilo
21 42 29.8 19 22 AAB51033 IgE peptide #11.
22 41 29.1 20 18 AAW24103 Canine immunoglobi
23 40 28.4 17 17 AAW08011 V3 peptide #8. Hu
24 39 27.7 9 21 AAY69599 Monoclonal antibody
25 39 27.7 9 21 AAY69600 Monoclonal antibody
26 39 27.7 9 21 AAY50891 Antibody 15A.2 bin
27 39 27.7 9 21 AAY50892 Antibody 15A.2 bin
28 39 27.7 14 21 AAY69579 Immunogenic peptid
29 38 27.0 12 21 AAB40058 Anti-hiL12 antibod
30 38 27.0 14 21 AAB26519 Human IgE C epsilo
31 37.5 26.6 16 21 AAB24095 Human apoptosis re
32 37.5 26.6 22 20 AAW97830 Human titin fibron
33 37 26.2 12 21 AAB40060 Anti-hiL12 antibod
34 37 26.2 12 21 AAB26516 Human IgE C epsilo
35 37 26.2 12 22 AAB51038 IgE peptide #16.
36 37 26.2 17 11 AAR02179 Peptide which bind
37 37 26.2 20 20 AAW74259 HJ loop peptide Cs
38 36 25.5 14 22 AAW00660 Human protein frag
39 36 25.5 18 19 AAW66675 HSV-2 glycoprotein
40 36 25.5 22 21 AAB21083 GDF-8 inhibitory p
41 35.5 25.2 18 22 AAB50754 Human CAMP-specifi
42 35 24.8 12 21 AAB40020 Anti-hiL12 antibod
43 35 24.8 12 21 AAB40059 Anti-hiL12 antibod
44 35 24.8 22 16 AAR72686 Peptide #3 having
45 34.5 24.5 16 15 AAR57918 Randomly generated
```

ALIGNMENTS

RESULT 1

AAAY79999
ID AAY79999 standard; Peptide; 25 AA.

XX AAY79999;

XX AC

XX DT 15-MAY-2000 (first entry)

XX DE Optimised IgE-CH3 domain antigen peptide for dog IgE.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;

XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Canis sp.

XX OS Synthetic.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX DR WPI; 2000-160578/14.

XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

XX PT for immunization against allergy

XX Claim 1: Page 99; 155pp; English.
 XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX Sequence 25 AA;

Query Match 100.0%; Score 141; DB 21; Length 25;
 Best Local Similarity 100.0%; Pred. No. 7.8e-15;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
 ID AAY79998
 DB 1 cgetyysrvthplpkdivrsiakc 25

RESULT 2
 AAY79998
 ID AAY79998 standard; Peptide; 25 AA.

XX AAY79998;
 XX 15-MAY-2000 (first entry)

DE Optimised IgE-CH3 domain antigen peptide for human IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Homo sapiens.
 OS Synthetic.

XX WO9967293-A1.
 PN 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -
 XX Claim 1: Page 21; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

Query Match 76.6%; Score 108; DB 21; Length 25;
 Best Local Similarity 72.0%; Pred. No. 8.8e-10;
 Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
 DB 1 cgetyysrvthplpralmrsttkc 25

RESULT 3
 AAY91212
 ID AAY91212 standard; peptide; 25 AA.

XX AAY91212;
 XX 22-MAY-2000 (first entry)

DE Modified human IgE CH3 domain, SEQ ID NO:92.

XX Promiscuous T-cell epitope; measles virus F protein; MVP;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;
 KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX Homo sapiens.
 OS Synthetic.

XX WO9966957-A2.
 PN 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13975.

XX 20-JUN-1998; 98US-0100412.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI; 2000-160564/14.

XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus -
 XX Example 6; Page 40; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesteryl ester transport

protein (CTPP) or HIV epitopes, but more generally against any pathogen, immunoreactive self-antigen or tumour antigen. The Th epitopes and peptide immunogens may be used for prevention and/or treatment of infections (HIV, foot-and-mouth disease or malaria); for cancer immunotherapy; for inhibition of the action of luteinising hormone releasing hormone (LHRH) for contraception, treatment of hormone-dependent cancer, prevention of boar taint in meat, and immunocastration; for promoting the growth of animals; or for treating allergies or arteriosclerosis. Incorporation of a promiscuous Th (functional in genetically diverse subjects) into an immunogen improves capacity to induce a strong T helper cell-mediated immune response, resulting in production of antibodies against a target antigen. Th can replace carrier proteins and pathogen-derived T helper epitopes. Sequence AAY91121 represents a promiscuous T helper epitope from the measles virus F (MVF) protein and sequences AAY91122-Y91142, AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope from hepatitis B virus (HBV) surface antigen, and sequences AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope. AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197 is the LHRH target antigenic peptide used in these LHRH antigenic peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic peptides comprising somatostatin and a Th epitope. Somatostatin immunogens may be used to promote growth in livestock. AAY91208 is a human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV infection of T cells. AAY90212 is a modified version of a human IgE (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3 antigenic peptides which may be used in the treatment of allergies. AAY91220 is a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent CTEP-derived peptides and AAY91232-Y91241 are immunogens comprising a CTEP peptide and a Th epitope which may be used to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1 vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory invasin protein epitope from Yersinia species, and hinge spacer peptide, both of which may optionally be used in the antigenic peptides of the invention.

Sequence 25 AA;

Query Match 76.6%; Score 108; DB 21; Length 25;
Best Local Similarity 72.0%; Pred. No. 8.8e-10;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||| ||||| :||| ||
Db 1 cgetygsrvthplpralmrsttkc 25

RESULT 4
AAY68602
ID AAY68602 standard; peptide; 25 AA.
XX
AC AAY68602;

XX
DT 05-MAY-2000 (first entry)
XX
DE Peptide sequence of the invention.
XX
KW Helper T cell epitope; peptide immunogen; LHRH;
KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.

XX Unidentified.
OS
XX WO9966952-A1.
PN
XX 29-DEC-1999.
PD
XX
PF 21-JUN-1999; 99WO-US13960.
XX
PR 20-JUN-1998; 98US-0100414.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
XX WPI; 2000-160562/14.
DR
XX New peptide immunogen containing luteinising hormone-releasing hormone antigen site and helper T cell epitope, for e.g. contraception and treatment of cancer.
PT
XX Disclosure; Page 92; 102pp; English.
XX
XX The specification describes peptide immunogens comprising a synthetic helper T cell (Th) epitope and a target antigen, luteinising hormone-releasing hormone (LHRH). The peptide immunogens cause induction of a specific immune response to LHRH which is involved in regulation of spermatogenesis, ovulation, oestrus, sexual development and secretion of sex hormones. Provision of a promiscuous T helper epitope (which is functional in genetically diverse subjects) provides optimum immunogenicity to the B cell epitopes of the target antigen and thus high antibody titres against the target antigen. The peptide immunogens of the invention are used to vaccinate against mammalian LHRH, for use as (reversible) contraceptive; control of hormone-dependent tumours (cancer of prostate or breast, also endometriosis); to prevent boar taint (and improve meat quality) and for immunocastration. The present sequence appears in the specification.

Sequence 25 AA;

Query Match 76.6%; Score 108; DB 21; Length 25;
Best Local Similarity 72.0%; Pred. No. 8.8e-10;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||| ||||| :||| ||
Db 1 cgetygsrvthplpralmrsttkc 25

RESULT 5
AAY80000
ID AAY80000 standard; Peptide; 25 AA.
XX
AC AAY80000;

XX
DT 15-MAY-2000 (first entry)
XX
DE Optimised IgE-CH3 domain antigen peptide for rat IgE.
XX
KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
OS Rattus sp.
OS Synthetic.
XX
XX WO9967293-A1.
PN
XX 29-DEC-1999.
PD
XX 21-JUN-1999; 99WO-US13959.

```

XX PR 20-JUN-1998; 98US-0100287.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY, Walfield AM;
XX PS WPI; 2000-160578/14.
XX CC
XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies
XX CC specific for a target effector site on the epsilon-heavy chain of IgE,
XX CC and so preventing triggering and activation of mast cells and basophils
XX CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
XX CC containing (I) are used for active immunisation against IgE-mediated
XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
XX CC dermatitis. Nucleic acids that encode these compounds are useful for
XX CC recombinant production of corresponding peptides or in DNA vaccines.
XX CC Conjugates of (I) that include a promiscuous T helper cell epitope
XX CC (functional in genetically diverse subjects), in addition to a B cell
XX CC target epitope, have increased immunogenicity and may include cyclic
XX CC constraints (disulfide bridge) to stabilise conformational features and
XX CC maximize cross-reactivity to the natural target. They induce safe
XX CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
XX CC acid sequences used in the exemplification of the present invention.
XX SQ Sequence 25 AA;

Query Match 74.5%; Score 105; DB 21; Length 25;
Best Local Similarity 76.0%; Pred. No. 2.5e-09;
Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
DB IIII I III III III III III

RESULT 6
AAW80077
ID AAY80077 standard; Peptide; 25 AA.
XX AC AAY80077;
XX DT 15-MAY-2000 (first entry)
XX DE Optimised IgE-CH3 domain antigen peptide for horse IgE.
XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;
XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX OS Equus caballus.
XX OS Synthetic.
XX PN WO9967293-A1.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-US13959.
XX PR 20-JUN-1998; 98US-0100287.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY, Walfield AM;

WPI; 2000-160578/14.
New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
for immunization against allergy -
Claim 1; Page 99; 155pp; English.
The present invention describes immunoglobulin E (IgE)-CH3 domain
antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
and anti-asthmatic properties. (I) induces polyclonal antibodies
specific for a target effector site on the epsilon-heavy chain of IgE,
and so preventing triggering and activation of mast cells and basophils
and downregulation of IgE synthesis. Conjugates, or fusion peptides,
containing (I) are used for active immunisation against IgE-mediated
allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
dermatitis. Nucleic acids that encode these compounds are useful for
recombinant production of corresponding peptides or in DNA vaccines.
Conjugates of (I) that include a promiscuous T helper cell epitope
(functional in genetically diverse subjects), in addition to a B cell
target epitope, have increased immunogenicity and may include cyclic
constraints (disulfide bridge) to stabilise conformational features and
maximize cross-reactivity to the natural target. They induce safe
(non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
acid sequences used in the exemplification of the present invention.
Sequence 25 AA;

Query Match 74.5%; Score 105; DB 21; Length 25;
Best Local Similarity 72.0%; Pred. No. 2.5e-09;
Matches 18; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
DB IIII I III III III III III
1 cgetykvshndpivrvrsiakc 25

RESULT 7
AAW24102
ID AAW24102 standard; peptide; 20 AA.
XX AC AAW24102;
XX DT 21-NOV-1997 (first entry)
XX DE Canine immunoglobulin E peptide 5.
XX KW Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.
XX OS Canis familiaris.
XX PN JP09169795-A.
XX PD 30-JUN-1997.
XX PF .22-DEC-1995; 95JP-0334381.
XX PR 22-DEC-1995; 95JP-0334381.
XX PA (HITB ) HITACHI CHEM CO LTD.
XX DR WPI; 1997-389423/36.
XX DR N-PSDB; AAT85651.
XX PT Canine immunoglobulin E peptide fragment and related DNA - useful
XX PT for the preparation of anti-canine immunoglobulin E antibody
XX PS Claim 2; Page 9; 12pp; Japanese.
XX CC AAW24098-106 are peptide fragments containing at least 5 continuous
XX CC amino acids of the partial canine immunoglobulin E (IgE) protein shown
XX CC in AAW24097. The peptides are used for the preparation of anti-canine

```



```

XX PD 12-OCT-1999.
XX XX
XX PF 21-APR-1994; 94US-0232539.
XX XX
XX PR 14-AUG-1991; 91US-0744768.
XX PR 07-JAN-1994; 94US-0178583.
XX XX
XX PA (GETH ) GENENTECH INC.
XX XX
XX PI Jardieu PM, Presta LG;
XX DR WPI; 1999-579941/49.
XX XX
XX PT Immunoglobulin E variants as peptide antagonists useful for raising and
PT screening anti-immunoglobulin E (IgE) antibodies, in the isolation and
PT purification of FcεpsilonRI receptor and in the treatment of allergic
PT diseases -
XX XX
XX PS Disclosure; Column 9; 37pp; English.
XX XX
XX CC The invention provides immunoglobulin E (IgE) antagonists comprising one
XX or more of the FcεpsilonRI receptor-binding determinant sites of human
XX IgE. The antagonists include IgE variants comprising an immunoglobulin
XX template and binding determinant sequences (bds) CDBds, EFBds and the
XX sequence shown in AAY42581. The CDBds (CD loop binding determinant
XX sequence) are selected from the sequences shown in AAY42567-Y42577 and
XX the EFBds (EF loop binding determinant sequence) are selected from
XX sequences shown in AAY42578-Y42580. The variants are useful in raising
XX and screening anti-IgE antibodies, in the isolation and purification of
XX FcεpsilonRI receptor and in the treatment and prophylaxis of allergic
XX diseases.
XX CC
XX SQ Sequence 23 AA;

Query Match 51.1%; Score 72; DB 20; Length 22;
Best Local Similarity 66.7%; Pred. No. 0.00024;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIV 19
Db 5 getyqcrvthplpralm 22

RESULT 11
AAY42586
ID AAY42586 standard; peptide; 24 AA.
XX
XX AC AAY42586;
XX XX
XX DT 10-JAN-2000 (first entry)
XX XX
XX DE IgE peptide antagonist.
XX XX
XX KW Immunoglobulin E; IgE; antagonist; FcεpsilonRI receptor; human; bds;
XX KW receptor-binding; binding determinant sequence; anti-IgE antibody;
XX KW allergic disease.
XX XX
XX OS Homo sapiens.
XX XX
XX PN US5965709-A.
XX XX
XX PD 12-OCT-1999.
XX XX
XX PF 21-APR-1994; 94US-0232539.
XX XX
XX PR 14-AUG-1991; 91US-0744768.
XX PR 07-JAN-1994; 94US-0178583.
XX XX
XX PA (GETH ) GENENTECH INC.
XX XX
XX PI Jardieu PM, Presta LG;
XX CC
XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX

```

```

XX DR WPI; 1999-579941/49.
XX XX
XX PT Immunoglobulin E variants as peptide antagonists useful for raising and
XX screening anti-immunoglobulin E (IgE) antibodies, in the isolation and
XX purification of FcεpsilonRI receptor and in the treatment of allergic
XX diseases -
XX XX
XX PS Disclosure; Column 9; 37pp; English.
XX XX
XX CC The invention provides immunoglobulin E (IgE) antagonists comprising one
XX or more of the FcεpsilonRI receptor-binding determinant sites of human
XX IgE. The antagonists include IgE variants comprising an immunoglobulin
XX template and binding determinant sequences (bds) CDBds, EFBds and the
XX sequence shown in AAY42581. The CDBds (CD loop binding determinant
XX sequence) are selected from the sequences shown in AAY42567-Y42577 and
XX the EFBds (EF loop binding determinant sequence) are selected from
XX sequences shown in AAY42578-Y42580. The variants are useful in raising
XX and screening anti-IgE antibodies, in the isolation and purification of
XX FcεpsilonRI receptor and in the treatment and prophylaxis of allergic
XX diseases.
XX CC
XX SQ Sequence 24 AA;

Query Match 51.1%; Score 72; DB 20; Length 24;
Best Local Similarity 66.7%; Pred. No. 0.00027;
Matches 12; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIV 19
Db 7 getyqcrvthplpralm 24

RESULT 12
AAY80001
ID AAY80001 standard; Peptide; 25 AA.
XX
XX AC AAY80001;
XX XX
XX DT 15-MAY-2000 (first entry)
XX XX
XX DE Optimised IgE-CH3 domain antigen peptide for mouse IgE.
XX XX
XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;
XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX XX
XX OS Mus sp.
XX OS Synthetic.
XX XX
XX PN WO9967293-A1.
XX XX
XX PD 29-DEC-1999.
XX XX
XX PF 21-JUN-1999; 99WO-US13959.
XX XX
XX PR 20-JUN-1998; 98US-0100287.
XX XX
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX XX
XX PI Wang CY, Walfield AM;
XX XX
XX DR WPI; 2000-160578/14.
XX XX
XX PF New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
XX for immunization against allergy -
XX XX
XX PS Claim 1; Page 100; 155pp; English.
XX XX
XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
XX

```


AC AAY50895;
XX
DT 24-FEB-2000 (first entry)
XX
DE Antibody 15A.2 human IgE binding epitope 1.
XX
KW Canine; allergy; antibody 15A.2; IgE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
XX
OS Synthetic.
XX
PN EP957111-A2.
XX
PD 17-NOV-1999.
XX
PF 09-APR-1999; 99EP-0107035.
XX
PR 09-APR-1998; 98US-0058331.
PR 30-MAR-1999; 99US-0281760.
XX
PA (IDEX-) IDEXX LAB INC.
XX
XX Lawton R, Mermer B, Francoeur G;
PI WPI; 2000-040833/04.
DR
XX
PT Binding proteins used for treatment or prophylaxis of canine allergy -
PS Disclosure; Fig 7; 30pp; English.
XX
CC This invention describes a novel binding protein which specifically
CC binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to IgE when the IgE is bound to mast cells. The peptide products of the
CC invention have anti-allergic activity. The antibodies bind to defined
CC epitopes on free or B-cell bound IgE molecules which have an important
CC role in allergic reaction. The specific binding proteins are used to
CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
CC invention.
XX
SQ Sequence 17 AA;

Query Match 42.6%; Score 60; DB 21; Length 17;
Best Local Similarity 71.4%; Pred. No. 0.012;
Matches 10; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 8 RVTHPHLPKDIVRS 21
| | | | | | | | : | |
Db 3 rvthphlpalmrs 16

Search completed: March 4, 2002, 13:10:08
Job time: 705 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:50:53 ; Search time 82.62 Seconds
(without alignments)
22.414 Million cell updates/sec

Title: US-09-701-623c-5
Perfect score: 140
Sequence: 1 CGEYQSRVTHPLRALMRSTTKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|----------|--------------------|
| 1 | 140 | 100.0 | 25 | AA199998 | Optimised IgE-CH3 |
| 2 | 140 | 100.0 | 25 | AA191212 | Modified human IgE |
| 3 | 140 | 100.0 | 25 | AA198602 | Peptide sequence o |
| 4 | 140 | 100.0 | 42 | AA180014 | IgE immunogenic pe |
| 5 | 140 | 100.0 | 42 | AA180013 | Modified MWF Th ep |
| 6 | 140 | 100.0 | 42 | AA191216 | Modified MWF Th ep |
| 7 | 140 | 100.0 | 42 | AA191217 | Modified MWF Th ep |
| 8 | 140 | 100.0 | 42 | AA191217 | Peptide sequence o |
| 9 | 140 | 100.0 | 45 | AA180007 | IgE-CH3 domain ant |
| 10 | 140 | 100.0 | 45 | AA191218 | Modified HBV surfa |
| 11 | 140 | 100.0 | 45 | AA198605 | Peptide sequence o |

| | | | | | | |
|----|-----|-------|-----|----|----------|-------------------------|
| 12 | 140 | 100.0 | 46 | 21 | AA180011 | IgE immunogenic pe |
| 13 | 140 | 100.0 | 46 | 21 | AA191213 | Modified MWF Th ep |
| 14 | 140 | 100.0 | 46 | 21 | AA191214 | Peptide sequence o |
| 15 | 140 | 100.0 | 46 | 21 | AA198603 | IgE immunogenic pe |
| 16 | 140 | 100.0 | 56 | 21 | AA180016 | IgE immunogenic pe |
| 17 | 140 | 100.0 | 59 | 21 | AA180010 | IgE immunogenic pe |
| 18 | 140 | 100.0 | 60 | 21 | AA180013 | IgE immunogenic pe |
| 19 | 140 | 100.0 | 60 | 21 | AA180015 | IgE immunogenic pe |
| 20 | 140 | 100.0 | 63 | 21 | AA180008 | IgE-CH3 domain ant |
| 21 | 140 | 100.0 | 63 | 21 | AA180012 | IgE immunogenic pe |
| 22 | 140 | 100.0 | 63 | 21 | AA191219 | Inv epitope/modifi |
| 23 | 140 | 100.0 | 63 | 21 | AA191219 | Peptide sequence o |
| 24 | 133 | 95.0 | 60 | 21 | AA180078 | IgE immunogenic pe |
| 25 | 117 | 83.6 | 106 | 20 | AA142620 | Human IgE Fc epsilon 3. |
| 26 | 117 | 83.6 | 110 | 14 | AA193304 | Variant IgE - muta |
| 27 | 117 | 83.6 | 110 | 14 | AA193314 | Variant IgE - muta |
| 28 | 117 | 83.6 | 110 | 14 | AA193315 | Variant IgE - muta |
| 29 | 117 | 83.6 | 110 | 14 | AA193318 | Variant IgE - muta |
| 30 | 117 | 83.6 | 110 | 14 | AA193319 | Variant IgE - muta |
| 31 | 117 | 83.6 | 110 | 14 | AA193320 | Variant IgE - muta |
| 32 | 117 | 83.6 | 110 | 14 | AA193321 | Variant IgE - muta |
| 33 | 117 | 83.6 | 110 | 14 | AA193322 | Variant IgE - muta |
| 34 | 117 | 83.6 | 110 | 14 | AA193327 | Variant IgE - muta |
| 35 | 117 | 83.6 | 110 | 14 | AA193328 | Variant IgE - muta |
| 36 | 117 | 83.6 | 110 | 14 | AA193329 | Variant IgE - muta |
| 37 | 117 | 83.6 | 110 | 14 | AA193330 | Variant IgE - muta |
| 38 | 117 | 83.6 | 110 | 14 | AA193331 | Variant IgE - muta |
| 39 | 117 | 83.6 | 110 | 14 | AA193332 | Variant IgE - muta |
| 40 | 117 | 83.6 | 110 | 14 | AA193333 | Variant IgE - muta |
| 41 | 117 | 83.6 | 110 | 14 | AA193334 | Variant IgE - muta |
| 42 | 117 | 83.6 | 110 | 14 | AA193335 | Variant IgE - muta |
| 43 | 117 | 83.6 | 110 | 14 | AA193336 | Variant IgE - muta |
| 44 | 117 | 83.6 | 110 | 14 | AA193337 | Variant IgE - muta |
| 45 | 117 | 83.6 | 110 | 14 | AA193338 | Variant IgE - muta |

ALIGNMENTS

RESULT 1
AA199998
ID AA199998 standard; Peptide; 25 AA.
XX
XX AA199998;
XX
XX
XX 15-MAY-2000 (first entry)
XX
XX Optimised IgE-CH3 domain antigen peptide for human IgE.
XX
XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX WO9967293-A1.
XX
XX
XX 29-DEC-1999
XX
XX 21-JUN-1999; 99WO-US13959.
XX
XX 20-JUN-1998; 98US-0100287.
XX
XX (UNBI-) UNITED BIOMEDICAL INC.
XX
XX Wang CY, Walfeld AM;
XX
XX WPI; 2000-160578/14.
XX
XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy

xx PS Claim 1; Page 21; 155pp; English.

xx CC The present invention describes immunoglobulin E (IgE)-CH3 domain

xx CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

xx CC and anti-asthmatic properties. (I) induces polyclonal antibodies

xx CC specific for a target effector site on the epsilon-heavy chain of IgE,

xx CC and so preventing triggering and activation of mast cells and basophils

xx CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

xx CC containing (I) are used for active immunisation against IgE-mediated

xx CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

xx CC dermatitis. Nucleic acids that encode these compounds are useful for

xx CC recombinant production of corresponding peptides or in DNA vaccines.

xx CC Conjugates of (I) that include a promiscuous T helper cell epitope

xx CC (functional in genetically diverse subjects), in addition to a B cell

xx CC target epitope, have increased immunogenicity and may include cyclic

xx CC constraints (disulfide bridge) to stabilise conformational features and

xx CC maximize cross-reactivity to the natural target. They induce safe

xx CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

xx CC acid sequences used in the exemplification of the present invention.

xx SQ Sequence 25 AA;

Query Match 100.0%; Score 140; DB 21; Length 25;

Best Local Similarity 100.0%; Pred. No. 4.8e-15;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGEYQSRVTHPLPALMRSTTKC 25

Db 1 cgetyqsrvtphlpralmrsttkc 25

RESULT 2

AAY91212

ID AAY91212 standard; peptide: 25 AA.

XX AC AAY91212;

XX DT 22-MAY-2000 (first entry)

XX DE Modified human IgE CH3 domain, SEQ ID NO:92.

XX KW Promiscuous T-cell epitope; measles virus F protein; MVF;

XX KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;

XX KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;

XX KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;

XX KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;

XX KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;

XX KW cholesterol ester transport protein; anti-arteriosclerotic.

OS Homo sapiens.

OS Synthetic.

XX WO9966957-A2.

XX PN 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13975.

XX PR 20-JUN-1998; 98US-0100412.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY;

XX WPI: 2000-160564/14.

XX New artificial T helper cell epitope and derived immunogens with target

XX PT antigenic site, for immunization against e.g. malaria, arteriosclerosis

XX PT or human immune deficiency virus

XX Example 6; Page 40; 129pp; English.

xx CC The invention relates to novel promiscuous T helper cell epitopes (Th),

xx CC and immunogenic peptides comprising the Th epitopes of the invention

xx CC along with B cell epitopes. The Th epitopes and peptide immunogens

xx CC containing them, are used to induce a T helper cell response,

xx CC specifically against Plasmodium falciparum, cholesterol ester transport

xx CC protein (CERP) or HIV epitopes, but more generally against any pathogen,

xx CC immunoreactive self-antigen or tumour antigen. The Th epitopes and

xx CC peptide immunogens may be used for prevention and/or treatment of

xx CC infections (HIV, foot-and-mouth disease or malaria); for cancer

xx CC immunotherapy; for inhibition of the action of luteinising hormone

xx CC releasing hormone (LHRH) for contraception, treatment of hormone-

xx CC dependent cancer, prevention of boar taint in meat, and

xx CC immunocastration; for promoting the growth of animals; or for

xx CC treating allergies or arteriosclerosis. Incorporation of a promiscuous

xx CC Th (functional in genetically diverse subjects) into an immunogen

xx CC improves capacity to induce a strong T helper cell-mediated immune

xx CC response, resulting in production of antibodies against a target

xx CC antigen. Th can replace carrier proteins and pathogen-derived T helper

xx CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope

xx CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,

xx CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the

xx CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope

xx CC from hepatitis B virus (HBV) surface antigen, and sequences

xx CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.

xx CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides

xx CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197

xx CC is the LHRH target antigenic peptide used in these LHRH antigenic

xx CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic

xx CC peptides comprising somatostatin and a Th epitope. Somatostatin

xx CC immunogens may be used to promote growth in livestock. AAY91208 is a

xx CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th

xx CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV

xx CC infection of T cells. AAY90212 is a modified version of a human IgE

xx CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3

xx CC antigenic peptides which may be used in the treatment of allergies.

xx CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)

xx CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th

xx CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target

xx CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th

xx CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent

xx CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a

xx CC CERP peptide and a Th epitope which may be used to prevent or treat

xx CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257

xx CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and

xx CC AAY91256-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell

xx CC epitope which may be used as a component in an anti-HIV-1 vaccine.

xx CC AAY91198 and AAY91199 are respectively an immunostimulatory Invasin

xx CC protein epitope from Yersinia species, and hinge spacer peptide, both of

xx CC which may optionally be used in the antigenic peptides of the

xx CC invention.

xx SQ Sequence 25 AA;

Query Match 100.0%; Score 140; DB 21; Length 25;

Best Local Similarity 100.0%; Pred. No. 4.8e-15;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGEYQSRVTHPLPALMRSTTKC 25

Db 1 cgetyqsrvtphlpralmrsttkc 25

RESULT 3

AAY68602

ID AAY68602 standard; peptide: 25 AA.

XX AC AAY68602;

XX DT 05-MAY-2000 (first entry)

XX DE Peptide sequence of the invention.

XX KW Helper T cell epitope; peptide immunogen; LHRH;
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
 XX OS Unidentified.
 XX WO9966952-A1.
 XX PD 29-DEC-1999.
 XX PF 21-JUN-1999; 99WO-US13960.
 XX PR 20-JUN-1998; 98US-0100414.
 XX PA (UNBI-) UNITED BIOMEDICAL INC.
 XX PI Wang CY;
 XX DR WPI; 2000-160562/14.
 XX PT New peptide immunogen containing luteinising hormone-releasing hormone
 PT antigen site and helper T cell epitope, for e.g. contraception and
 PT treatment of cancer
 XX PS Disclosure; Page 92; 102pp; English.
 CC The specification describes peptide immunogens comprising a
 CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
 CC hormone-releasing hormone (LHRH). The peptide immunogens cause
 CC induction of a specific immune response to LHRH which is involved in
 CC regulation of spermatogenesis, ovulation, oestrus, sexual development
 CC and secretion of sex hormones. Provision of a promiscuous T helper
 CC epitope (which is functional in genetically diverse subjects) provides
 CC optimum immunogenicity to the B cell epitopes of the target antigen and
 CC thus high antibody titres against the target antigen. The peptide
 CC immunogens of the invention are used to vaccinate against mammalian LHRH,
 CC for use as (reversible) contraceptive; control of hormone-dependent
 CC tumours (cancer of prostate or breast, also endometriosis); to prevent
 CC boar taint (and improve meat quality) and for immunocastration. The
 CC present sequence appears in the specification.
 XX SQ Sequence 25 AA;
 Query Match 100.0%; Score 140; DB 21; Length 25;
 Best Local Similarity 100.0%; Pred. No. 4.8e-15;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGETYQSRVTHPLPALMRSTTKC 25
 ID AAY80014
 XX AAY80014 standard; Peptide; 42 AA.
 AC AAY80014;
 XX 15-MAY-2000 (first entry)
 XX IgE immunogenic peptide conjugate SEQ ID NO:21.
 XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX OS Unidentified.

PN WO9967293-A1.
 XX 29-DEC-1999.
 XX PF 21-JUN-1999; 99WO-US13959.
 XX PR 20-JUN-1998; 98US-0100287.
 XX PA (UNBI-) UNITED BIOMEDICAL INC.
 XX PI Wang CY, Walfield AM;
 XX DR WPI; 2000-160578/14.
 XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy
 XX PS Claim 14; Page 76; 155pp; English.
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX SQ Sequence 42 AA;
 Query Match 100.0%; Score 140; DB 21; Length 42;
 Best Local Similarity 100.0%; Pred. No. 8.7e-15;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGETYQSRVTHPLPALMRSTTKC 25
 ID AAY91215
 XX AAY91215 standard; peptide; 42 AA.
 AC AAY91215;
 XX 22-MAY-2000 (first entry)
 XX Modified MVF Th epitope/IgE CH3 domain, SEQ ID NO:95.
 XX Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;
 KW cholesterol ester transport protein; anti-arteriosclerotic.
 XX Chimeric - Measles virus.
 OS Chimeric - Homo sapiens.
 XX WO9966957-A2.
 XX 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-USI3975.
XX PR 20-JUN-1998; 98US-0100412.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PT Wang CY;
XX PS WPI: 2000-160564/14.
XX DR New artificial T helper cell epitope and derived immunogens with target
XX PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
XX PT or human immune deficiency virus
XX PS Example 6; Page 98-99; 129pp; English.
XX CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesterol ester transport
CC protein (CTP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration; for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (WVF) protein and sequences AAY91122-Y91142,
CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC WVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a
CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AAY90212 is a modified version of a human IgE
CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AAY91224-Y91225 comprise the CS antigen and an WVF Th
CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
CC CTP-derived peptides and AAY91232-Y91241 are immunogens comprising a
CC CTP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
CC protein epitope from Yersinia species, and hinge spacer peptide, both of
CC which may optionally be used in the antigenic peptides of the
CC invention.
XX SQ Sequence 42 AA;

Query Match 100.0%; Score 140; DB 21; Length 42;
Best Local Similarity 100.0%; Pred. No. 8.7e-15;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CGETYQSRVTHPLPRALMRSTTKC 25
Db 18 cgetyqsrvtphlpralmrsttkc 42
RESULT 6
AAY91216
ID AAY91216 standard; peptide; 42 AA.
XX AC AAY91216;
XX DT 22-MAY-2000 (first entry)
XX DE Modified WVF Th epitope/IgE CH3 domain, SEQ ID NO:96.
XX KW Promiscuous T-cell epitope; measles virus F protein; WVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CTP;
KW cholesterol ester transport protein; anti-arteriosclerotic.
XX OS Chimeric - Measles virus.
OS Chimeric - Homo sapiens.
XX PN WO9966957-A2.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-USI3975.
XX PR 20-JUN-1998; 98US-0100412.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY;
XX WPI: 2000-160564/14.
XX PT New artificial T helper cell epitope and derived immunogens with target
XX PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
XX PT or human immune deficiency virus
XX PS Example 6; Page 99; 129pp; English.
XX CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesterol ester transport
CC protein (CTP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration; for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (WVF) protein and sequences AAY91122-Y91142,
CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC WVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a
CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AAY90212 is a modified version of a human IgE
CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AAY91224-Y91225 comprise the CS antigen and an WVF Th
CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
CC CTP-derived peptides and AAY91232-Y91241 are immunogens comprising a
CC CTP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
CC protein epitope from Yersinia species, and hinge spacer peptide, both of
CC which may optionally be used in the antigenic peptides of the
CC invention.
XX SQ Sequence 42 AA;

CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

XX Sequence 42 AA;

Query Match 100.0%; Score 140; DB 21; Length 42;
 Best Local Similarity 100.0%; Pred. No. 8.7e-15;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPHLPALMRSTTKC 25
 Db 18 cgetyqsrvtphlpalmrsttkc 42

RESULT 7
 AAY91217
 ID AAY91217 standard; peptide; 42 AA.

XX AAY91217;

XX 22-MAY-2000 (first entry)

XX Modified MVF Th epitope/IgE CH3 domain, SEQ ID NO:97.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
 KW cholesterol ester transport protein; anti-arteriosclerotic.

XX Chimeric - Measles virus.
 OS Chimeric - Homo sapiens.

XX W09966957-A2.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13975.

XX 20-JUN-1998; 98US-0100412.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI; 2000-160564/14.

XX

PT New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 or human immune deficiency virus
 XX Example 6; Page 99; 129pp; English.
 XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesterol ester transport
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone-
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration); for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.

XX Sequence 42 AA;

Query Match 100.0%; Score 140; DB 21; Length 42;
 Best Local Similarity 100.0%; Pred. No. 8.7e-15;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPHLPALMRSTTKC 25
 Db 18 cgetyqsrvtphlpalmrsttkc 42

RESULT 8
 AAY68604
 ID AAY68604 standard; peptide; 42 AA.
 XX

AC AAY68604;
 XX
 DT 05-MAY-2000 (first entry)
 XX
 DE Peptide sequence of the invention.
 XX
 KW Helper T cell epitope; peptide immunogen; LHRH;
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1
 FT /label- Ile, Met, Leu
 FT Misc-difference 2
 FT /label- Ser, Thr
 FT Misc-difference 5
 FT /label- Lys, Arg
 FT Misc-difference 6
 FT /label- Gly, Thr
 FT Misc-difference 10
 FT /label- His, Thr
 FT Misc-difference 11
 FT /label- Lys, Arg
 FT Misc-difference 12
 FT /label- Ile, Met, Leu
 FT Misc-difference 14
 FT /label- Gly, Thr
 FT Misc-difference 15
 FT /label- Ile, Met, Val
 XX
 PN WO9966952-A1.
 XX
 XX 29-DEC-1999.
 XX
 XX 21-JUN-1999; 99WO-US13960.
 XX
 XX 20-JUN-1998; 98US-0100414.
 XX
 XX (UNBI-) UNITED BIOMEDICAL INC.
 XX
 XX Wang CY;
 XX WPI: 2000-160562/14.
 XX
 XX New peptide immunogen containing luteinising hormone-releasing hormone
 XX antigen site and helper T cell epitope, for e.g. contraception and
 XX treatment of cancer -
 XX
 XX Disclosure; Page 95; 102pp; English.
 XX
 XX The specification describes peptide immunogens comprising a
 XX synthetic helper T cell (Th) epitope and a target antigen, luteinising
 XX hormone-releasing hormone (LHRH). The peptide immunogens cause
 XX induction of a specific immune response to LHRH which is involved in
 XX regulation of spermatogenesis, ovulation, oestrus, sexual development
 XX and secretion of sex hormones. Provision of a promiscuous T helper
 XX epitope (which is functional in genetically diverse subjects) provides
 XX optimum immunogenicity to the B cell epitopes of the target antigen and
 XX thus high antibody titres against the target antigen. The peptide
 XX immunogens of the invention are used to vaccinate against mammalian LHRH,
 XX for use as (reversible) contraceptive; control of hormone-dependent
 XX tumours (cancer of prostate or breast, also endometriosis); to prevent
 XX boar taint (and improve meat quality) and for immunocastration. The
 XX present sequence appears in the specification.
 XX
 XX Sequence 42 AA;

Query Match 100.0%; Score 140; DB 21; Length 42;

Best Local Similarity 100.0%; Pred. No. 8.7e-15;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGETYQSRVTHPLPRALMRSTTKC 25
 DB 18 cgetyqsrvtphlpralmrsttkc 42
 RESULT 9
 AAY80007
 ID AAY80007 standard; Peptide; 45 AA.
 XX
 XX AAY80007;
 XX
 XX 15-MAY-2000 (first entry)
 DT
 XX
 DE IgE-CH3 domain antigen peptide antigenic site 15b SEQ ID NO:14.
 XX
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Unidentified.
 XX
 PN WO9967293-A1.
 XX
 PD 29-DEC-1999.
 XX
 XX 21-JUN-1999; 99WO-US13959.
 XX
 XX 20-JUN-1998; 98US-0100287.
 XX
 XX (UNBI-) UNITED BIOMEDICAL INC.
 XX
 XX Wang CY, Walfield AM;
 XX WPI: 2000-160578/14.
 XX
 XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 XX for immunization against allergy -
 XX
 XX Claim 14; Page 75; 155pp; English.
 XX
 XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 XX antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 XX and anti-asthmatic properties. (I) induces polyclonal antibodies
 XX specific for a target effector site on the epsilon-heavy chain of IgE,
 XX and so preventing triggering and activation of mast cells and basophils
 XX and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 XX containing (I) are used for active immunisation against IgE-mediated
 XX allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 XX dermatitis. Nucleic acids that encode these compounds are useful for
 XX recombinant production of corresponding peptides or in DNA vaccines.
 XX Conjugates of (I) that include a promiscuous T helper cell epitope
 XX (functional in genetically diverse subjects), in addition to a B cell
 XX target epitope, have increased immunogenicity and may include cyclic
 XX constraints (disulfide bridge) to stabilise conformational features and
 XX maximize cross-reactivity to the natural target. They induce safe
 XX (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 XX acid sequences used in the exemplification of the present invention.
 XX
 XX Sequence 45 AA;

Query Match 100.0%; Score 140; DB 21; Length 45;
 Best Local Similarity 100.0%; Pred. No. 9.4e-15;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYQSRVTHPLPRALMRSTTKC 25
 DB 21 cgetyqsrvtphlpralmrsttkc 45

RESULT 10
 AAY91218
 ID AAY91218 standard; peptide; 45 AA.
 XX
 AC AAY91218;
 XX
 DT 22-MAY-2000 (first entry)
 DE
 DE Modified HBV surface Ag/IgE CH3 domain, SEQ ID NO:98.
 XX
 KW Promiscuous T-cell epitope; measles virus F protein; WVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumporoite; antimalarial; CEMP;
 KW cholesterol ester transport protein; anti-arteriosclerotic.
 XX
 OS Chimeric - Hepatitis B virus.
 OS Chimeric - Homo sapiens.
 XX
 PN WO9966957-A2.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13975.
 XX
 PR 20-JUN-1998; 98US-0100412.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY;
 XX
 XX WPI; 2000-160564/14.
 XX
 PT New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus
 XX
 XX Example 6; Page 100; 129pp; English.
 PS
 CC The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesterol ester transport
 CC protein (CEP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration; for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (WVF) protein and sequences AAY91122-Y91142,
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC WVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th

CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumporoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an WVF Th
 CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CEP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CEP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
 CC protein epitope from *Yersinia* species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.
 XX
 SQ Sequence 45 AA;

Query Match 100.0%; Score 140; DB 21; Length 45;
 Best Local Similarity 100.0%; Pred. No. 9.4e-15;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGETYQSRVTHPLPALMRSTTKC 25
 |||||
 Db 21 cgetyqsrvtphlplalmrsttkc 45

RESULT 11
 AAY68605
 ID AAY68605 standard; peptide; 45 AA.
 XX
 AC AAY68605;
 XX
 DT 05-MAY-2000 (first entry)
 DE
 DE Peptide sequence of the invention.
 XX
 KW Helper T cell epitope; peptide immunogen; LHRH;
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
 XX
 OS Unidentified.
 XX WO9966952-A1.
 XX
 PD 29-DEC-1999.
 PF 21-JUN-1999; 99WO-US13960.
 XX
 PR 20-JUN-1998; 98US-0100414.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY;
 XX
 XX WPI; 2000-160562/14.
 XX
 PT New peptide immunogen containing luteinising hormone-releasing hormone
 PT antigen site and helper T cell epitope, for e.g. contraception and
 PT treatment of cancer
 XX
 PS Disclosure; Page 95; 102pp; English.
 CC The specification describes peptide immunogens comprising a
 CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
 CC hormone-releasing hormone (LHRH). The peptide immunogens cause

induction of a specific immune response to LHRH which is involved in regulation of spermatogenesis, ovulation, oestrus, sexual development and secretion of sex hormones. Provision of a promiscuous T helper epitope (which is functional in genetically diverse subjects) provides optimum immunogenicity to the B cell epitopes of the target antigen and thus high antibody titres against the target antigen. The peptide immunogens of the invention are used to vaccinate against mammalian LHRH, for use as (reversible) contraceptive; control of hormone-dependent tumours (cancer of prostate or breast, also endometriosis); to prevent boar taint (and improve meat quality) and for immunocastration. The present sequence appears in the specification.

| Sequence | 45 AA; |
|----------|--------|
| SQ | |

Query Match 100.0%; Score 140; DB 21; Length 45;
Best Local Similarity 100.0%; Pred. No. 9.4e-15;
Matches 25; Conservative 0; Mismatches 0; Indels

Qy 1 CGETYQSFVTHPHLPALMRSTTKC 25
 |||||
 Db 21 caetvqsrvtphlpalmrsttkc 45

CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.

Sequence 46 AA;

Query Match 100.0%; Score 140; DB 21; Length 46;
Best Local Similarity 100.0%; Pred. No. 9.6e-15;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGETYQSRVTHPHLPALMRSTTKC 25
+++++
Db 22 cgetygsrvthphlpalmrsttkc 46

QY 1 CGETYQSRVTHPHLPALMRSTTKC 25
Db 22 cgetyqsrvtphlpalmrsttkc 46

RESULT 15

AAV68603
ID AAV68603 standard; peptide; 46 AA.

AC AAV68603;

XX 05-MAY-2000 (first entry)

XX Peptide sequence of the invention.

XX Helper T cell epitope; peptide immunogen; LHRH;
KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.

XX Unidentified.

XX Key Location/Qualifiers

XX Misc-difference 4 /label= Ser, Thr
FT Misc-difference 7 /label= Lys, Arg
FT Misc-difference 8 /label= Gly, Thr
FT Misc-difference 12 /label= His, Thr
FT Misc-difference 13 /label= Lys, Arg
FT Misc-difference 16 /label= Gly, Thr

XX WO9966952-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13960.

XX 20-JUN-1998; 98US-0100414.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI; 2000-160562/14.

XX New peptide immunogen containing luteinising hormone-releasing hormone
PT antigen site and helper T cell epitope, for e.g. contraception and
PT treatment of cancer

XX Disclosures; Page 93; 102pp; English.

XX The specification describes peptide immunogens comprising a
CC synthetic helper T cell (Th) epitope and a target antigen, luteinising
CC hormone-releasing hormone (LHRH). The peptide immunogens cause
CC induction of a specific immune response to LHRH which is involved in
CC regulation of spermatogenesis, ovulation, oestrus, sexual development
CC and secretion of sex hormones. Provision of a promiscuous T helper
CC epitope which is functional in genetically diverse subjects provides
CC optimum immunogenicity to the B cell epitopes of the target antigen and
CC thus high antibody titres against the target antigen. The peptide
CC immunogens of the invention are used to vaccinate against mammalian LHRH,
CC for use as (reversible) contraceptive; control of hormone-dependent
CC tumours (cancer of prostate or breast, also endometriosis); to prevent
CC boar taint (and improve meat quality) and for immunocastration. The
CC present sequence appears in the specification.

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:55:50 ; Search time 82.62 seconds
(without alignments)
22.414 Million cell updates/sec

Title: US-09-701-623C-6

Perfect score: 141
Sequence: 1 CGETYYSRVTHPLPKDIVRSIAKC 25

Scoring table: BLCSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | ID | Description |
|------------|-------|-------|--------|-------------|--------------------|
| 1 | 141 | 100.0 | 25 | 21 AAY79999 | Optimised IgE-CH3 |
| 2 | 141 | 100.0 | 45 | 21 AAY80019 | IgE immunogenic pe |
| 3 | 141 | 100.0 | 45 | 21 AAY80083 | IgE immunogenic pe |
| 4 | 141 | 100.0 | 46 | 21 AAY80020 | IgE immunogenic pe |
| 5 | 141 | 100.0 | 57 | 21 AAY80081 | IgE immunogenic pe |
| 6 | 141 | 100.0 | 62 | 21 AAY80080 | IgE immunogenic pe |
| 7 | 141 | 100.0 | 63 | 21 AAY80084 | IgE immunogenic pe |
| 8 | 118 | 83.7 | 124 | 18 AAW24097 | Partial canine imm |
| 9 | 118 | 83.7 | 312 | 21 AAY79995 | Dog immunoglobulin |
| 10 | 118 | 83.7 | 341 | 21 AAB06208 | Immunogenic peptid |
| 11 | 118 | 83.7 | 417 | 18 AAW23067 | Canine IgE heavy c |

| | | | | | |
|----|-----|------|-----|-------------|--------------------|
| 12 | 118 | 83.7 | 426 | 17 AAR97753 | Canine IgE. Canis |
| 13 | 108 | 76.6 | 25 | 21 AAY79998 | Optimised IgE-CH3 |
| 14 | 108 | 76.6 | 25 | 21 AAY91212 | Modified human IgE |
| 15 | 108 | 76.6 | 25 | 21 AAY68602 | Peptide sequence o |
| 16 | 108 | 76.6 | 42 | 21 AAY80014 | IgE immunogenic pe |
| 17 | 108 | 76.6 | 42 | 21 AAY91215 | Modified MVF Th ep |
| 18 | 108 | 76.6 | 42 | 21 AAY91216 | Modified MVF Th ep |
| 19 | 108 | 76.6 | 42 | 21 AAY91217 | Modified MVF Th ep |
| 20 | 108 | 76.6 | 42 | 21 AAY68604 | Peptide sequence o |
| 21 | 108 | 76.6 | 45 | 21 AAY80007 | IgE-CH3 domain ant |
| 22 | 108 | 76.6 | 45 | 21 AAY91218 | Modified HBV surfa |
| 23 | 108 | 76.6 | 45 | 21 AAY68605 | Peptide sequence o |
| 24 | 108 | 76.6 | 46 | 21 AAY80011 | IgE immunogenic pe |
| 25 | 108 | 76.6 | 46 | 21 AAY91213 | Modified MVF Th ep |
| 26 | 108 | 76.6 | 46 | 21 AAY91214 | Modified MVF Th ep |
| 27 | 108 | 76.6 | 46 | 21 AAY68603 | Peptide sequence o |
| 28 | 108 | 76.6 | 56 | 21 AAY80016 | IgE immunogenic pe |
| 29 | 108 | 76.6 | 59 | 21 AAY80010 | IgE immunogenic pe |
| 30 | 108 | 76.6 | 60 | 21 AAY80013 | IgE immunogenic pe |
| 31 | 108 | 76.6 | 60 | 21 AAY80015 | IgE immunogenic pe |
| 32 | 108 | 76.6 | 63 | 21 AAY80008 | IgE-CH3 domain ant |
| 33 | 108 | 76.6 | 63 | 21 AAY80012 | IgE immunogenic pe |
| 34 | 108 | 76.6 | 63 | 21 AAY91219 | Inv epitope/modifi |
| 35 | 108 | 76.6 | 63 | 21 AAY68606 | Peptide sequence o |
| 36 | 105 | 74.5 | 25 | 21 AAY80000 | Optimised IgE-CH3 |
| 37 | 105 | 74.5 | 25 | 21 AAY80077 | Optimised IgE-CH3 |
| 38 | 101 | 71.6 | 60 | 21 AAY80078 | IgE immunogenic pe |
| 39 | 95 | 67.4 | 20 | 18 AAW24102 | Canine immunoglobi |
| 40 | 91 | 64.5 | 345 | 21 AAB06207 | Immunogenic peptid |
| 41 | 86 | 61.0 | 17 | 21 AAY50893 | Antibody 15A.2 bin |
| 42 | 86 | 61.0 | 17 | 21 AAY50894 | Antibody 15A.2 can |
| 43 | 86 | 61.0 | 110 | 14 AAR33324 | Variant IgE - muta |
| 44 | 85 | 60.3 | 106 | 20 AAY42620 | Human IgE Fcpsiilo |
| 45 | 85 | 60.3 | 110 | 14 AAR33304 | IgE Fc epsilon 3. |

ALIGNMENTS

RESULT 1

AAV79999

ID AAY79999 standard; Peptide; 25 AA.

XX AC AAY79999;

XX AC AAY79999;

DT 15-MAY-2000 (first entry)

XX 15-MAY-2000 (first entry)

DE Optimised IgE-CH3 domain antigen peptide for dog IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Canis sp.

OS Synthetic.

XX WO967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy

XX Claim 1; Page 59; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils

CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

CC containing (I) are used for active immunisation against IgE-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell

CC target epitope, have increased immunogenicity and may include cyclic

CC constraints (disulfide bridge) to stabilise conformational features and

CC maximize cross-reactivity to the natural target. They induce safe

CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

CC acid sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

QY 1 CGETYSRVTHPLPKDIVRSIAKC 25

Db 1 cgetysrvthplpkdivrsiakc 25

RESULT 2

AA80019

ID AAY80019 standard; Peptide; 45 AA.

XX AAY80019;

AC

XX 15-MAY-2000 (first entry)

DT

XX IgE immunogenic peptide conjugate SEQ ID NO:26.

DE

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

OS

XX WO9967293-A1.

PN

XX 29-DEC-1995.

PD

XX 21-JUN-1999; 99WO-US13959.

XX

XX 20-JUN-1998; 98US-0100287.

XX

XX (UNBI-) UNITED BIOMEDICAL INC.

PA

XX Wang CY, Walfield AM;

PI

XX WPI; 2000-160578/14.

DR

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy

PS

XX Claim 14; Page 76; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils

CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

CC containing (I) are used for active immunisation against IgE-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell

CC target epitope, have increased immunogenicity and may include cyclic

CC constraints (disulfide bridge) to stabilise conformational features and

CC maximize cross-reactivity to the natural target. They induce safe

CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

CC acid sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

QY 1 CGETYSRVTHPLPKDIVRSIAKC 25

Db 1 cgetysrvthplpkdivrsiakc 25

RESULT 3

AA80083

ID AAY80083 standard; Peptide; 45 AA.

XX AAY80083;

AC

XX 15-MAY-2000 (first entry)

DT

XX IgE immunogenic peptide conjugate SEQ ID NO:90.

DE

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

OS

XX WO9967293-A1.

PN

XX 29-DEC-1999.

PD

XX 21-JUN-1999; 99WO-US13959.

XX

XX 20-JUN-1998; 98US-0100287.

XX

XX (UNBI-) UNITED BIOMEDICAL INC.

PA

XX Wang CY, Walfield AM;

PI

XX WPI; 2000-160578/14.

DR

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy

PS

XX Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils

CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

CC containing (I) are used for active immunisation against IgE-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY7994 to AAY8084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 45 AA;

Query Match 100.0%; Score 141; DB 21; Length 45;
 Best Local Similarity 100.0%; Pred. No. 1.6e-14;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
 Db 21 cgetyysrvthplpkdivrsiakc 45
 |||||

RESULT 4
 AAY80020
 ID AAY80020 standard; Peptide; 46 AA.
 AC AAY80020;
 XX
 XX DT 15-MAY-2000 (first entry)
 XX IgE immunogenic peptide conjugate SEQ ID NO:27.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.
 XX WO9967293-A1.

XX 29-DEC-1999.
 XX 21-JUN-1999; 99WO-US13959.
 XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -

XX Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY7994 to AAY8084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX

SQ Sequence 46 AA;

Query Match 100.0%; Score 141; DB 21; Length 46;
 Best Local Similarity 100.0%; Pred. No. 1.6e-14;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
 Db 22 cgetyysrvthplpkdivrsiakc 46
 |||||

RESULT 5
 AAY80081
 ID AAY80081 standard; Peptide; 57 AA.

XX AAY80081;

XX 15-MAY-2000 (first entry)

XX IgE immunogenic peptide conjugate SEQ ID NO:88.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -

XX Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY7994 to AAY8084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 57 AA;

Query Match 100.0%; Score 141; DB 21; Length 57;
 Best Local Similarity 100.0%; Pred. No. 2e-14;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
 Db 33 cgetyysrvthplpkdivrsiakc 57

RESULT 6

AAV80080
 ID AAV80080 standard; Peptide: 62 AA.

AC AAV80080;

DT 15-MAY-2000 (first entry)

DE IGE immunogenic peptide conjugate SEQ ID NO:87.

OS Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

PN W09967293-A1.

PD 29-DEC-1999.

PF 21-JUN-1999; 99WO-US13959.

PR 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

PA Wang CY, Walfield AM;

PI WPT; 2000-160578/14.

DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -

PS Claim 14; Page 77; 155pp; English.

CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAV79994 to AAV80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX Sequence 62 AA;

Query Match 100.0%; Score 141; DB 21; Length 62;
 Best Local Similarity 100.0%; Pred. No. 2.3e-14;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25

Db 38 cgetyysrvthplpkdivrsiakc 62

RESULT 7

AAV80084

ID AAV80084 standard; Peptide: 63 AA.

AC AAV80084;

DT 15-MAY-2000 (first entry)

DE IGE immunogenic peptide conjugate SEQ ID NO:91.

OS Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

PN W09967293-A1.

PD 29-DEC-1999.

PF 21-JUN-1999; 99WO-US13959.

PR 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

PA Wang CY, Walfield AM;

PI WPT; 2000-160578/14.

DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -

PS Claim 14; Page 77; 155pp; English.

CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAV79994 to AAV80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX Sequence 63 AA;

Query Match 100.0%; Score 141; DB 21; Length 63;
 Best Local Similarity 100.0%; Pred. No. 2.3e-14;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25

Db 39 cgetyysrvthplpkdivrsiakc 63

RESULT 8

AAW24097

ID AAW24097 standard; peptide: 124 AA.

AC AAW24097;

DT 21-NOV-1997 (first entry)

DE Partial canine immunoglobulin E protein.

XX Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.
 XX
 OS Canis familiaris.
 XX
 PN JP09169795-A.
 XX
 PD 30-JUN-1997.
 XX
 PF 22-DEC-1995; 95JP-0334381.
 XX
 XX 22-DEC-1995; 95JP-0334381.
 XX
 XX (HITB) HITACHI CHEM CO LTD.
 XX
 DR WPI; 1997-389423/36.
 DR N-PSDB; AAT85646.
 XX
 XX Canine immunoglobulin E peptide fragment and related DNA - useful
 PT for the preparation of anti-canine immunoglobulin E antibody
 XX
 PS Claim 1; Page 8; 12pp; Japanese.
 XX
 CC This is a partial canine immunoglobulin E (IgE) protein. Peptide
 CC fragments (AAW24098-106) containing at least five continuous amino acids
 CC of this sequence are used for the preparation of anti-canine IgE
 CC antibody. The anti-canine IgE antibody can be used for the diagnosis of
 CC canine allergies.
 XX
 SQ Sequence 124 AA;
 Query Match 83.7%; Score 118; DB 18; Length 124;
 Best Local Similarity 95.7%; Pred. No. 1.7e-10;
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GETYYSRVTHPLPKDIVRSIAK 24
 Db 57 getycrvrthplpkdivrsiak 79
 RESULT 9
 AAY79995
 ID AAY79995 standard; Protein; 312 AA.
 XX
 AC AAY79995;
 XX
 DT 15-MAY-2000 (first entry)
 XX
 DE Dog immunoglobulin E epsilon heavy chain SEQ ID NO:2.
 XX
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Canis sp.
 XX
 PN WO9967293-A1.
 XX
 PD 29-DEC-1999.
 XX
 XX 21-JUN-1999; 99WO-US13959.
 XX
 XX 20-JUN-1998; 98US-0100287.
 XX
 XX (UNBI-) UNITED BIOMEDICAL INC.
 PA
 PI Wang CY, Walfield AM;
 XX
 DR WPI; 2000-160578/14.
 XX
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy -
 XX
 XX Example 1; Page 66-68; 155pp; English.
 XX
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 312 AA;
 Query Match 83.7%; Score 118; DB 21; Length 312;
 Best Local Similarity 95.7%; Pred. No. 4.9e-10;
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GETYYSRVTHPLPKDIVRSIAK 24
 Db 187 getycrvrthplpkdivrsiak 209
 RESULT 10
 AAB06208
 ID AAB06208 standard; protein; 341 AA.
 XX
 AC AAB06208;
 XX
 DT 22-NOV-2000 (first entry)
 XX
 DE Immunogenic peptide consisting of opossum CH2, dog CH3 and opossum CH4.
 XX
 KW Dog; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
 KW asthma; eczema; immunogenic peptide.
 XX
 OS Chimeric - Didelphis virginiana.
 OS Chimeric - Canis sp.
 XX
 PN WO200025722-A2.
 XX
 PD 11-MAY-2000.
 XX
 PF 21-OCT-1999; 99WO-SE01896.
 XX
 PR 02-NOV-1998; 98US-0106652.
 PR 22-SEP-1999; 99US-0401636.
 XX
 XX (RESI-) RESISTENTIA PHARM AB.
 PA
 PI Hellman LT;
 XX
 XX WPI; 2000-365342/31.
 XX
 XX Immunogenic polypeptides useful for preventing the harmful effects of
 PT immunoglobulin E in mammals -
 XX
 XX Disclosure; Fig 2; 50pp; English.
 XX
 CC The present sequence is an immunogenic peptide consisting of the heavy
 CC chain constant regions 2 and 4 of the opossum IgE and the heavy chain
 CC constant region 3 from the dog. It was shown to cause a stronger

CC polyclonal anti-self IgE response than peptides consisting of the same
CC regions from one mammal. Immunogenic peptides, particularly those
CC consisting of different heavy chain constant regions, can be used for
CC vaccination in humans, against bacterial and viral infections and
CC allergies, such as asthma, fur, pollen and food allergies and eczema.
XX
SQ Sequence 341 AA;

Query Match 83.7%; Score 118; DB 21; Length 341;
Best Local Similarity 95.7%; Pred. No. 5.4e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYSRVTHPLPKDIVRSIAK 24
||||| ||||||| ||||||| |||||
DB 206 getyyrcrvthphlpkdivrsiak 228

RESULT 11
AAW23067
ID AAW23067 standard; Protein; 417 AA.

XX AC AAW23067;

DT 19-FEB-1998 (first entry)

DE Canine IgE heavy chain constant region (exon 1-4 product).

XX IgE; immunoglobulin; antibody; heavy chain constant region;
KW allergy; hypersensitivity; therapy; dog; antisense;
KW immunomodulation.

XX Canis familiaris.

XX Key Location/Qualifiers

FT Misc-difference 55 /note= "encoded by ACC"
FT Misc-difference 56 /note= "encoded by TAC"
FT Misc-difference 67 /note= "encoded by GCC"
FT Misc-difference 83 /note= "encoded by NNT"
FT Misc-difference 174 /note= "encoded by GGN"
FT Misc-difference 175 /note= "encoded by NNG"
FT Misc-difference 176 /note= "encoded by TGN"
FT Misc-difference 203 /note= "encoded by TCC"
FT Misc-difference 204 /note= "encoded by GAC"

XX W09730156-A2.

XX 21-AUG-1997.

XX 14-FEB-1997; 97WO-US02322.

XX 14-FEB-1996; 96US-0601197.

XX (IDEX-) IDEXX LAB INC.

XX Harris RA, Mermer B, Siefring AE;

XX WPI; 1997-425031/39.

XX N-PSDB; AAT79278.

XX Isolated canine IgE heavy chain constant region DNA - useful to
XX develop products for treatment of canine allergies and for
XX immunomodulation in dogs

PS Disclosure; Page 35-39; 59pp; English.
XX This polypeptide is encoded by exons 1-4 (see AAT79278) of canine
CC IgE heavy chain constant region (epsilon) genomic DNA. Another
CC polypeptide, comprising the exon 5 and 6 product, is given in
CC AAW23068. Recombinant peptides encoded by exons 1-6 can be
CC produced in eukaryotic or prokaryotic cells. Such peptides,
CC and antibodies raised against them, are used in methods to treat
CC the manifestation of allergy in dogs, e.g. to treatment Type I
CC immediate hypersensitivity, and for immunomodulation.
XX
SQ Sequence 417 AA;

Query Match 83.7%; Score 118; DB 18; Length 417;
Best Local Similarity 95.7%; Pred. No. 6.9e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYSRVTHPLPKDIVRSIAK 24
||||| ||||||| ||||||| |||||
DB 284 getyyrcrvthphlpkdivrsiak 306

RESULT 12
AAR97753
ID AAR97753 standard; Protein; 426 AA.

XX AC AAR97753;

DT 28-AUG-1996 (first entry)

DE Canine IgE.

XX IgE; immunoglobulin; antibody; vaccine; allergy; hypersensitivity.

XX Canis familiaris.

XX W09614867-A1.

XX 23-MAY-1996.

XX 03-NOV-1995; 95WO-US13795.

XX 09-NOV-1994; 94US-0336891.

XX 09-NOV-1994; 94US-0336583.

XX (MERI) MERCK & CO INC.

XX Hollis GF, Patel MD;

XX WPI; 1996-277321/28.

XX N-PSDB; AAT29824.

XX New DNA encoding canine IgE and IgA - useful in vaccines, antisense
XX therapy, assays, drug screening, etc.

XX Claim 11; Page 29-30; 49pp; English.

XX The canine IgE amino acid sequence (AAR97753) was deduced from
CC an isolated gene (AAT29824) obt'd. from a canine liver DNA library.
CC The cloning of the IgE gene allows prodn. of large quantities of
CC recombinant IgE using bacterial, yeast, mammalian, insect or
CC viral systems. The IgE can be used in drug development (e.g.
CC small molecule screening, assay development and anti-IgE
CC antibody generation). Fragments of IgE can be used in vaccines
CC or to prevent IgE-mediated hypersensitivity. The new sequence
CC information permits targeted modulation of IgE-mediated immune
CC responses.

XX Sequence 426 AA;

Query Match 83.7%; Score 118; DB 17; Length 426;

Best Local Similarity 95.7%; Pred. No. 7e-10; Mismatches 0; Indels 1; Gaps 0; Matches 22; Conservative 0;

QY 2 GETYSRVTHPLPKDIVRSIAK 24
||||| ||||||| |||||||
Db 289 getycrvthplpkdivrsiak 311

RESULT 13
AA79998
ID AAY79998 standard; Peptide; 25 AA.

XX AC AAY79998;
XX DT 15-MAY-2000 (first entry)
XX DE Optimised IgE-CH3 domain antigen peptide for human IgE.
XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Homo sapiens.
XX OS Synthetic.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX DR WPI; 2000-160578/14.

XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy

XX PS Claim 1; Page 21; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 25 AA;

Query Match 76.6%; Score 108; DB 21; Length 25;
Best Local Similarity 72.0%; Pred. No. 8.8e-10;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPLPKDIVRSIAK 25
||||| ||||||| ||| :||| ||
Db 1 cgetysrvthplpkdivrsiatkc 25

RESULT 14

AA91212

ID AAY91212 standard; peptide; 25 AA.

XX AC AAY91212;

XX DT 22-MAY-2000 (first entry)

XX DE Modified human IgE CH3 domain, SEQ ID NO:92.

XX KW Promiscuous T-cell epitope; measles virus F protein; MFP;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;
KW cholesterol ester transport protein; anti-arteriosclerotic.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO9966957-A2.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13975.

XX PR 20-JUN-1998; 98US-0100412.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY;

XX DR WPI; 2000-160564/14.

XX PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus

XX PS Example 6; Page 40; 129pp; English.

XX CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesterol ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone-
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration; for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC MFP Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a

CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
 CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
 CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
 CC epitope, and may be used in a malaria vaccine. AAY91228-Y91231 represent
 CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
 CC CERP peptide and a Th epitope which may be used to prevent or treat
 CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
 CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
 CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
 CC epitope which may be used as a component in an anti-HIV-1 vaccine.
 CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
 CC protein epitope from Yersinia species, and hinge spacer peptide, both of
 CC which may optionally be used in the antigenic peptides of the
 CC invention.
 XX
 SQ Sequence 25 AA;

Query Match 76.6%; Score 108; DB 21; Length 25;
 Best Local Similarity 72.0%; Pred. No. 8.8e-10;
 Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGETYYSRVTHPLPKDIVRSIAKC 25
 ||||| ||||| ||||| :||| ||
 Db 1 cgetyqsrvtphplpralmrsttkc 25

RESULT 15
 AAY68602
 ID AAY68602 standard; peptide; 25 AA.
 XX
 AC AAY68602;
 XX
 XX 05-MAY-2000 (first entry)
 XX
 DE Peptide sequence of the invention.
 KW Helper T cell epitope; peptide immunogen; LHRH;
 KW luteinising hormone-releasing hormone; spermatogenesis; ovulation;
 KW oestrus; sexual development; sex hormone; promiscuous T helper epitope;
 KW vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
 KW breast cancer; endometriosis; boar taint; meat quality; immunocastration.
 XX
 OS Unidentified.
 XX
 PN WO9966952-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13960.
 XX
 PR 20-JUN-1998; 98US-0100414.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PT Wang CY;
 XX
 PT WPI; 2000-160562/14.
 DR
 XX
 PT New peptide immunogen containing luteinising hormone-releasing hormone
 PT antigen site and helper T cell epitope, for e.g. contraception and
 PT treatment of cancer
 XX
 XX Disclosure; Page 92; 102pp; English.
 XX
 CC The specification describes peptide immunogens comprising a
 CC synthetic helper T cell (Th) epitope and a target antigen, luteinising

CC hormone-releasing hormone (LHRH). The peptide immunogens cause
 CC induction of a specific immune response to LHRH which is involved in
 CC regulation of spermatogenesis, ovulation, oestrus, sexual development
 CC and secretion of sex hormones. Provision of a promiscuous T helper
 CC epitope (which is functional in genetically diverse subjects) provides
 CC optimum immunogenicity to the B cell epitopes of the target antigen and
 CC thus high antibody titres against the target antigen. The peptide
 CC immunogens of the invention are used to vaccinate against mammalian LHRH,
 CC for use as (reversible) contraceptive; control of hormone-dependent
 CC tumours (cancer of prostate or breast, also endometriosis); to prevent
 CC boar taint (and improve meat quality) and for immunocastration. The
 CC present sequence appears in the specification.
 XX
 SQ Sequence 25 AA;

Query Match 76.6%; Score 108; DB 21; Length 25;
 Best Local Similarity 72.0%; Pred. No. 8.8e-10;
 Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGETYYSRVTHPLPKDIVRSIAKC 25
 ||||| ||||| ||||| :||| ||
 Db 1 cgetyqsrvtphplpralmrsttkc 25

Search completed: March 4, 2002, 12:55:50
 Job time: 297 sec

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:55:50 ; Search time 82.62 Seconds
(without alignments)
22.414 Million cell updates/sec

Title: US-09-701-623C-7

Perfect score: 145

Sequence: 1 CGEGYQSRVDHPFKPIVRITK 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 145 | 100.0 | 25 | 21 | AAV80000 |
| 2 | 122 | 84.1 | 313 | 21 | AAV79996 |
| 3 | 122 | 84.1 | 340 | 21 | AA803643 |
| 4 | 122 | 84.1 | 341 | 21 | AA806206 |
| 5 | 121 | 83.4 | 45 | 21 | AAV80018 |
| 6 | 121 | 83.4 | 46 | 21 | AAV80017 |
| 7 | 106 | 73.1 | 25 | 21 | AAV80001 |
| 8 | 105 | 72.4 | 25 | 21 | AAV79999 |
| 9 | 105 | 72.4 | 45 | 21 | AAV80019 |
| 10 | 105 | 72.4 | 45 | 21 | AAV80083 |
| 11 | 105 | 72.4 | 46 | 21 | AAV80020 |
| | | | | | Optimised IgE-CH3 |
| | | | | | Rat immunoglobulin |
| | | | | | Rat IgE heavy chain |
| | | | | | Immunogenic peptid |
| | | | | | IgE immunogenic pe |
| | | | | | IgE immunogenic pe |
| | | | | | Optimised IgE-CH3 |
| | | | | | Optimised IgE-CH3 |
| | | | | | IgE immunogenic pe |
| | | | | | IgE immunogenic pe |

| | | | | | | |
|----|-----|------|-----|----|----------|--------------------|
| 12 | 105 | 72.4 | 57 | 21 | AAV80081 | IgE immunogenic pe |
| 13 | 105 | 72.4 | 62 | 21 | AAV80080 | IgE immunogenic pe |
| 14 | 105 | 72.4 | 63 | 21 | AAV80084 | Optimised IgE-CH3 |
| 15 | 103 | 71.0 | 25 | 21 | AAV79998 | Modified human IgE |
| 16 | 103 | 71.0 | 25 | 21 | AAV91212 | Peptide sequence o |
| 17 | 103 | 71.0 | 25 | 21 | AAV68602 | IgE immunogenic pe |
| 18 | 103 | 71.0 | 42 | 21 | AAV80014 | Modified MWF Th ep |
| 19 | 103 | 71.0 | 42 | 21 | AAV91215 | Modified MWF Th ep |
| 20 | 103 | 71.0 | 42 | 21 | AAV91216 | Modified MWF Th ep |
| 21 | 103 | 71.0 | 42 | 21 | AAV91217 | Peptide sequence o |
| 22 | 103 | 71.0 | 42 | 21 | AAV68604 | IgE-CH3 domain ant |
| 23 | 103 | 71.0 | 45 | 21 | AAV80007 | Modified HBV surfa |
| 24 | 103 | 71.0 | 45 | 21 | AAV91218 | Peptide sequence o |
| 25 | 103 | 71.0 | 45 | 21 | AAV68605 | IgE immunogenic pe |
| 26 | 103 | 71.0 | 46 | 21 | AAV80011 | Modified MWF Th ep |
| 27 | 103 | 71.0 | 46 | 21 | AAV91213 | Modified MWF Th ep |
| 28 | 103 | 71.0 | 46 | 21 | AAV91214 | Peptide sequence o |
| 29 | 103 | 71.0 | 46 | 21 | AAV68603 | IgE immunogenic pe |
| 30 | 103 | 71.0 | 56 | 21 | AAV80016 | IgE immunogenic pe |
| 31 | 103 | 71.0 | 59 | 21 | AAV80010 | IgE immunogenic pe |
| 32 | 103 | 71.0 | 60 | 21 | AAV80013 | IgE immunogenic pe |
| 33 | 103 | 71.0 | 60 | 21 | AAV80015 | IgE immunogenic pe |
| 34 | 103 | 71.0 | 63 | 21 | AAV80008 | IgE-CH3 domain ant |
| 35 | 103 | 71.0 | 63 | 21 | AAV80012 | IgE immunogenic pe |
| 36 | 103 | 71.0 | 63 | 21 | AAV91219 | Inv epitope/modifi |
| 37 | 103 | 71.0 | 63 | 21 | AAV68606 | Peptide sequence o |
| 38 | 96 | 66.2 | 60 | 21 | AAV80078 | IgE immunogenic pe |
| 39 | 90 | 62.1 | 313 | 21 | AAV79997 | Mouse immunoglobul |
| 40 | 89 | 61.4 | 25 | 21 | AAV80077 | Optimised IgE-CH3 |
| 41 | 85 | 58.6 | 561 | 20 | AAV17415 | Mouse immunoglobul |
| 42 | 84 | 57.9 | 110 | 14 | AAK33317 | Variant IgE - muta |
| 43 | 82 | 56.6 | 124 | 18 | AAK24097 | Partial canine imm |
| 44 | 82 | 56.6 | 312 | 21 | AAV79995 | Dog immunoglobulin |
| 45 | 82 | 56.6 | 341 | 21 | AA806208 | Immunogenic peptid |

ALIGNMENTS

RESULT 1

AAV80000
ID: AAV80000 standard; Peptide: 25 AA.

XX AAV80000;

XX AC

XX 15-MAY-2000 (first entry)

XX Optimised IgE-CH3 domain antigen peptide for rat IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Rattus sp.

OS Synthetic.

XX WO9804293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy

XX PS Claim 1; Page 59; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils

CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

CC containing (I) are used for active immunisation against IgE-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell

CC target epitope, have increased immunogenicity and may include cyclic

CC constraints (disulfide bridge) to stabilise conformational features and

CC maximize cross-reactivity to the natural target. They induce safe

CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 25 AA;

Query Match 100.0%; Score 145; DB 21; Length 25;

Best Local Similarity 100.0%; Pred. No. 4.7e-15;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITK 25

DB 1 CGEGYQSRVDHPHPKPIVRSITK 25

||||| ||||||| ||||||| |||||||

RESULT 2

AAY79996

ID AAY79996 standard; Protein; 313 AA.

XX AC AAY79996;

XX DT 15-MAY-2000 (first entry)

XX DE Rat Immunoglobulin E epsilon heavy chain SEQ ID NO:3.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Rattus sp.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfeld AM;

XX DR WPI; 2000-160578/14.

XX PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PS for immunization against allergy -

XX PS Example 1; Page 66-68; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils

CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

CC containing (I) are used for active immunisation against IgE-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell

CC target epitope, have increased immunogenicity and may include cyclic

CC constraints (disulfide bridge) to stabilise conformational features and

CC maximize cross-reactivity to the natural target. They induce safe

CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 313 AA;

Query Match 84.1%; Score 122; DB 21; Length 313;

Best Local Similarity 95.7%; Pred. No. 2.2e-10;

Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24

DB 190 GEGYQSRVDHPHPKPIVRSITK 212

||||| ||||||| ||||||| |||||||

RESULT 3

AAB03643

ID AAB03643 standard; protein; 340 AA.

XX AC AAB03643;

XX DT 22-NOV-2000 (first entry)

XX DE Rat IgE heavy chain constant regions 2, 3 and 4.

XX KW Rat; immunoglobulin E; IgE; vaccination; infection; allergy;

KW asthma; eczema; immunogenic peptide.

XX OS Rattus sp.

XX PN WO200025722-A2.

XX PD 11-MAY-2000.

XX PF 21-OCT-1999; 99WO-SE01896.

XX PR 02-NOV-1998; 98US-0106652.

XX PR 22-SEP-1999; 99US-0401636.

XX PA (RESI-) RESISTENTIA PHARM AB.

XX PI Hellman LT;

XX DR WPI; 2000-365342/31.

XX PT Immunogenic polypeptides useful for preventing the harmful effects of

XX PT immunoglobulin E in mammals -

XX PS Disclosure; Fig 1; 50pp; English.

XX CC The present sequence is an immunogenic peptide consisting of the

CC heavy chain constant regions 2, 3 and 4 of the rat IgE. It was used to

CC construct a number of immunogenic peptides which consisted of regions of

CC IgE from different mammals, which appear to cause a stronger polyclonal

CC anti-self IgE response than peptides consisting of the same regions from

CC one mammal. Immunogenic peptides, particularly those consisting of

CC different heavy chain constant regions, can be used for vaccination

CC in humans, against bacterial and viral infections and allergies, such

CC as asthma, fur, pollen and food allergies and eczema.

XX SQ Sequence 340 AA;

Query Match 84.1%; Score 122; DB 21; Length 340;
 Best Local Similarity 95.7%; Pred. No. 2.4e-10;
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
 ||||| ||||| ||||| ||||| |||||
 Db 195 gegyqcrvdhphfhpkipvrsitk 217

RESULT 4
 AAB06206
 ID AAB06206 standard; protein; 341 AA.
 XX
 AC AAB06206;
 XX
 DT 22-NOV-2000 (first entry)
 XX
 DE Immunogenic peptide consisting of opossum CH2, rat CH3 and opossum CH4.
 XX
 KW Rat; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
 KW asthma; eczema; immunogenic peptide.
 XX
 OS Chimeric - Didelphis virginiana.
 OS Chimeric - Rattus sp.
 XX
 PN WO200025722-A2.
 XX
 PD 11-MAY-2000.
 XX
 PF 21-OCT-1999; 99WO-S801896.
 XX
 PR 02-NOV-1998; 98US-0106652.
 PR 22-SEP-1999; 99US-0401636.
 XX
 PA (RESI-) RESISTENTIA PHARM AB.
 XX
 PI Hellman LT;
 XX
 DR WPI; 2000-365342/31.
 XX
 PT Immunogenic polypeptides useful for preventing the harmful effects of
 PT Immunoglobulin E in mammals -
 XX
 PS Disclosure; Fig 2; 50pp; English.
 CC
 CC The present sequence is an immunogenic peptide consisting of the heavy
 CC chain constant regions 2 and 4 of the opossum IgE and the heavy chain
 CC constant region 3 from the rat. It was shown to cause a stronger
 CC polyclonal anti-self IgE response than peptides consisting of the same
 CC regions from one mammal. Immunogenic peptides, particularly those
 CC consisting of different heavy chain constant regions, can be used for
 CC vaccination in humans, against bacterial and viral infections and
 CC allergies, such as asthma, fur, pollen and food allergies and eczema.
 XX
 SQ Sequence 341 AA;

Query Match 84.1%; Score 122; DB 21; Length 341;
 Best Local Similarity 95.7%; Pred. No. 2.4e-10;
 Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
 ||||| ||||| ||||| ||||| |||||
 Db 206 gegyqcrvdhphfhpkipvrsitk 228

RESULT 5
 AAY80018
 ID AAY80018 standard; Peptide; 45 AA.
 XX
 AC AAY80018;
 XX

DT 15-MAY-2000 (first entry)
 XX
 DE IgE immunogenic peptide conjugate SEQ ID NO:25.
 XX
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Unidentified.
 XX
 PN WO9967293-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13959.
 XX
 PR 20-JUN-1998; 98US-0100287.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY, Walfield AM;
 XX
 DR WPI; 2000-160578/14.
 XX
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -
 XX
 PS Claim 14; Page 76; 155pp; English.
 XX
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY7994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 45 AA;

Query Match 83.4%; Score 121; DB 21; Length 45;
 Best Local Similarity 88.0%; Pred. No. 3.6e-11;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITK 25
 ||||| ||||| ||||| ||||| |||||
 Db 21 cgygyqsrvdhphfhpkipvrsitkc 45

RESULT 6
 AAY80017
 ID AAY80017 standard; Peptide; 46 AA.
 XX
 AC AAY80017;
 XX
 DT 15-MAY-2000 (first entry)
 XX
 DE IgE immunogenic peptide conjugate SEQ ID NO:24.
 XX
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX Unidentified.
 OS
 XX
 PN WO9967293-A1.
 XX
 XX 29-DEC-1999.
 PD
 XX
 XX 21-JUN-1999; 99WO-US13959.
 PF
 XX
 PR 20-JUN-1998; 98US-0100287.
 XX
 XX (UNBI-) UNITED BIOMEDICAL INC.
 PA
 XX Wang CY, Walfield AM;
 PI
 XX WPI: 2000-160578/14.
 DR
 XX
 XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy
 PT
 XX
 PS Claim 14; Page 76; 155pp; English.
 XX
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 46 AA;
 Query Match 83.4%; Score 121; DB 21; Length 46;
 Best Local Similarity 88.0%; Pred. No. 3.7e-11;
 Matches 22; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
 DB 22 CGYGRQSLVDHPDPKPIVRSITKC 46
 RESULT 7
 AAY80001
 ID AAY80001 standard; Peptide; 25 AA.
 XX
 AC AAY80001;
 XX
 XX
 DT 15-MAY-2000 (first entry)
 XX
 DE Optimised IgE-CH3 domain antigen peptide for mouse IgE.
 XX
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Mus sp.
 OS Synthetic.
 XX
 XX WO9967293-A1.
 PN
 XX

PD 29-DEC-1999.
 XX
 XX 21-JUN-1999; 99WO-US13959.
 XX
 XX 20-JUN-1998; 98US-0100287.
 XX
 XX (UNBI-) UNITED BIOMEDICAL INC.
 PA
 XX Wang CY, Walfield AM;
 PI
 XX WPI: 2000-160578/14.
 DR
 XX
 XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy
 PT
 XX
 PS Claim 1; Page 100; 155pp; English.
 XX
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;
 Query Match 73.1%; Score 106; DB 21; Length 25;
 Best Local Similarity 80.0%; Pred. No. 3.4e-09;
 Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
 DB 1 CGYGRQSLVDHPDPKPIVRSITKC 25
 RESULT 8
 AAY79999
 ID AAY79999 standard; Peptide; 25 AA.
 XX
 AC AAY79999;
 XX
 XX
 DT 15-MAY-2000 (first entry)
 XX
 DE Optimised IgE-CH3 domain antigen peptide for dog IgE.
 XX
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Canis sp.
 OS Synthetic.
 XX
 XX WO9967293-A1.
 PN
 XX 29-DEC-1999.
 PD
 XX
 XX 21-JUN-1999; 99WO-US13959.
 PF
 XX
 XX 20-JUN-1998; 98US-0100287.
 XX
 XX

PA (UNBI-) UNITED BIOMEDICAL INC.
 XX Wang CY, Walfield AM;
 XX WPI; 2000-160578/14.
 DR
 XX
 XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy
 PT
 XX
 XX Claim 1; Page 99; 155pp; English.
 XX
 XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 XX Sequence 25 AA;
 SQ

Query Match 72.4%; Score 105; DB 21; Length 25;
 Best Local Similarity 76.0%; Pred. No. 4.8e-09;
 Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
 III I III III I IIIII II
 Db 1 cgetysrvthpklpdivrslakc 25

RESULT 9
 AAY80019
 ID AAY80019 standard; Peptide; 45 AA.
 XX
 AC AAY80019;
 XX
 DT 15-MAY-2000 (first entry)
 XX
 DE IgE immunogenic peptide conjugate SEQ ID NO:26.
 XX
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Unidentified.
 XX
 XX WO9967293-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13959.
 XX
 PR 20-JUN-1998; 98US-0100287.
 XX
 XX (UNBI-) UNITED BIOMEDICAL INC.
 XX
 XX Wang CY, Walfield AM;
 XX WPI; 2000-160578/14.
 DR
 XX
 XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy
 XX
 XX Claim 14; Page 76; 155pp; English.
 XX
 XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 XX Sequence 45 AA;
 SQ

Query Match 72.4%; Score 105; DB 21; Length 45;
 Best Local Similarity 76.0%; Pred. No. 9.2e-09;
 Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
 III I III III I IIIII II
 Db 21 cgetysrvthpklpdivrslakc 45

RESULT 10
 AAY80083
 ID AAY80083 standard; Peptide; 45 AA.
 XX
 AC AAY80083;
 XX
 DT 15-MAY-2000 (first entry)
 XX
 DE IgE immunogenic peptide conjugate SEQ ID NO:90.
 XX
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Unidentified.
 XX
 XX WO9967293-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13959.
 XX
 PR 20-JUN-1998; 98US-0100287.
 XX
 XX (UNBI-) UNITED BIOMEDICAL INC.
 XX
 XX Wang CY, Walfield AM;
 XX WPI; 2000-160578/14.
 DR
 XX
 XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy
 PT
 XX
 XX Claim 14; Page 77; 155pp; English.
 XX
 XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

SQ Sequence 45 AA;

Query Match 72.4%; Score 105; DB 21; Length 45;
 Best Local Similarity 76.0%; Pred. No. 9.2e-09;
 Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGEGYQSRVDHPHPKPIVRSITKC 25
 ||| | ||| ||| ||| ||| |||
 Db 21 cgetysrvthphlpkdivrsiack 45

RESULT 11

AAY80020
 ID AAY80020 standard; Peptide: 46 AA.

XX AC AAY80020;

DT 15-MAY-2000 (first entry)

XX IgE immunogenic peptide conjugate SEQ ID NO:27.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

XX WO9967293-A1.

PD 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -

XX Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

SQ Sequence 46 AA;

Query Match 72.4%; Score 105; DB 21; Length 46;
 Best Local Similarity 76.0%; Pred. No. 9.4e-09;
 Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CGEGYQSRVDHPHPKPIVRSITKC 25
 ||| | ||| ||| ||| ||| |||
 Db 22 cgetysrvthphlpkdivrsiack 46

RESULT 12

AAY80081
 ID AAY80081 standard; Peptide: 57 AA.

XX AC AAY80081;

DT 15-MAY-2000 (first entry)

XX IgE immunogenic peptide conjugate SEQ ID NO:88.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

XX WO9967293-A1.

PD 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -

XX Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

XX
SQ

Sequence 57 AA;

Query Match 72.4%; Score 105; DB 21; Length 57;

Best Local Similarity 76.0%; Pred. No. 1.2e-08;

Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Db 33 cgetysrvthphlpkdivrsiack 57

RESULT 13

AAAY80080

ID AAY80080 standard; Peptide; 62 AA.

XX AC AAY80080;

XX DT 15-MAY-2000 (first entry)

XX IgE immunogenic peptide conjugate SEQ ID NO:87.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Unidentified.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy

XX Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 62 AA;

Query Match

Best Local Similarity 72.4%; Score 105; DB 21; Length 62;

Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Db 38 cgetysrvthphlpkdivrsiack 62

RESULT 14

AAAY80084

ID AAY80084 standard; Peptide; 63 AA.

XX AC AAY80084;

XX DT 15-MAY-2000 (first entry)

XX IgE immunogenic peptide conjugate SEQ ID NO:91.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Unidentified.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy

XX Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 63 AA;

Query Match

Best Local Similarity 72.4%; Score 105; DB 21; Length 63;

Matches 19; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Db 39 cgetysrvthphlpkdivrsiack 63

RESULT 15

AAV79998
 ID AAV79998 standard; Peptide; 25 AA.
 XX
 AC AAY79998;
 XX
 DT 15-MAY-2000 (first entry)
 XX
 DE Optimised IgE-CH3 domain antigen peptide for human IgE.
 XX
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO9967293-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13959.
 XX
 PR 20-JUN-1998; 98US-0100287.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY, Walfield AM;
 XX
 DR WPI; 2000-163578/14.
 XX
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -
 XX
 PS Claim 1; Page 21; 155pp; English.
 XX
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAV79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;

Query Match 71.0%; Score 103; DB 21; Length 25;
 Best Local Similarity 68.0%; Pred. No. 9.6e-09;
 Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 CGEGYQSRVDHPHPKPIVRSITKC 25
 ||| ||||| ||| | : ||| |||
 Db 1 cgetyqsrvtphplpalmrsttkc 25

Search completed: March 4, 2002, 12:55:50
 Job time: 297 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March. 4, 2002, 12:55:50 ; Search time 82.62 Seconds
(without alignments)
22.414 Million cell updates/sec

Title: US-09-701-623C-8
Perfect score: 140
Sequence: 1 CGYQSVDRPDKPKIVRSITLC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101:*

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 140 | 100.0 | 25 | 21 | AA1980001 |
| 2 | 129 | 92.1 | 45 | 21 | AA1980018 |
| 3 | 129 | 92.1 | 46 | 21 | AA1980017 |
| 4 | 116 | 82.9 | 561 | 20 | AA17415 |
| 5 | 113 | 80.7 | 313 | 21 | AA179997 |
| 6 | 106 | 75.7 | 25 | 21 | AA1980000 |
| 7 | 85 | 60.7 | 313 | 21 | AA179996 |
| 8 | 85 | 60.7 | 340 | 21 | AA1980018 |
| 9 | 85 | 60.7 | 341 | 21 | AA1980017 |
| 10 | 74 | 52.9 | 25 | 21 | AA1980077 |
| 11 | 66 | 47.1 | 25 | 21 | AA179999 |

| | | | | | | |
|----|----|------|-----|----|-----------|--------------------|
| 12 | 66 | 47.1 | 45 | 21 | AA1980019 | IgE immunogenic pe |
| 13 | 66 | 47.1 | 45 | 21 | AA1980083 | IgE immunogenic pe |
| 14 | 66 | 47.1 | 46 | 21 | AA1980020 | IgE immunogenic pe |
| 15 | 66 | 47.1 | 57 | 21 | AA1980081 | IgE immunogenic pe |
| 16 | 66 | 47.1 | 62 | 21 | AA1980080 | IgE immunogenic pe |
| 17 | 66 | 47.1 | 63 | 21 | AA1980084 | IgE immunogenic pe |
| 18 | 64 | 45.7 | 25 | 21 | AA179998 | Optimised IgE-CH3 |
| 19 | 64 | 45.7 | 25 | 21 | AA191212 | Modified human IgE |
| 20 | 64 | 45.7 | 25 | 21 | AA198602 | Peptide sequence o |
| 21 | 64 | 45.7 | 25 | 21 | AA1980014 | IgE immunogenic pe |
| 22 | 64 | 45.7 | 42 | 21 | AA191215 | Modified MvF Th ep |
| 23 | 64 | 45.7 | 42 | 21 | AA191216 | Modified MvF Th ep |
| 24 | 64 | 45.7 | 42 | 21 | AA191217 | Modified MvF Th ep |
| 25 | 64 | 45.7 | 42 | 21 | AA198604 | Peptide sequence o |
| 26 | 64 | 45.7 | 45 | 21 | AA1980007 | IgE-CH3 domain ant |
| 27 | 64 | 45.7 | 45 | 21 | AA191218 | Modified HBV surfa |
| 28 | 64 | 45.7 | 45 | 21 | AA198605 | Peptide sequence o |
| 29 | 64 | 45.7 | 46 | 21 | AA1980011 | IgE immunogenic pe |
| 30 | 64 | 45.7 | 46 | 21 | AA191213 | Modified MvF Th ep |
| 31 | 64 | 45.7 | 46 | 21 | AA191214 | Modified MvF Th ep |
| 32 | 64 | 45.7 | 46 | 21 | AA198603 | Peptide sequence o |
| 33 | 64 | 45.7 | 56 | 21 | AA1980016 | IgE immunogenic pe |
| 34 | 64 | 45.7 | 59 | 21 | AA1980010 | IgE immunogenic pe |
| 35 | 64 | 45.7 | 60 | 21 | AA1980013 | IgE immunogenic pe |
| 36 | 64 | 45.7 | 60 | 21 | AA1980015 | IgE immunogenic pe |
| 37 | 64 | 45.7 | 63 | 21 | AA1980008 | IgE-CH3 domain ant |
| 38 | 64 | 45.7 | 63 | 21 | AA1980012 | IgE immunogenic pe |
| 39 | 64 | 45.7 | 63 | 21 | AA191219 | Inv epitope/modifi |
| 40 | 64 | 45.7 | 63 | 21 | AA198606 | Peptide sequence o |
| 41 | 63 | 45.0 | 60 | 21 | AA1980078 | IgE immunogenic pe |
| 42 | 58 | 41.4 | 345 | 21 | AA1986207 | Immunogenic peptid |
| 43 | 54 | 38.6 | 15 | 21 | AA1980899 | Antibody 15A.2 mur |
| 44 | 53 | 37.9 | 17 | 21 | AA1980898 | Antibody 15A.2 swi |
| 45 | 50 | 35.7 | 344 | 22 | AA1980826 | Micromonospora eve |

ALIGNMENTS

RESULT 1

AA1980001

ID AA1980001 standard; Peptide: 25 AA.

XX

AC AA1980001;

XX

DT 15-MAY-2000 (first entry)

XX

DE Optimised IgE-CH3 domain antigen peptide for mouse IgE.

XX

KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX

OS Mus sp.

OS Synthetic.

XX

PN WO996293-A1.

XX

PD 29-DEC-1999.

XX

PF 21-JUN-1999; 99WO-US13959.

XX

PR 20-JUN-1998; 98US-0100287.

XX

PA (UNBI-) UNITED BIOMEDICAL INC.

XX

PI Wang CV, Walfield AM;

XX

DR WPI; 2000-160578/14.

XX

PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy

XX Claim 1; Page 100; 155pp; English.

PS The present invention describes immunoglobulin E (IgE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils

CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

CC containing (I) are used for active immunisation against IgE-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell

CC target epitope, have increased immunogenicity and may include cyclic

CC constraints (disulfide bridge) to stabilise conformational features and

CC maximize cross-reactivity to the natural target. They induce safe

CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

CC acid sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

5Q

Query Match 100.0%; Score 140; DB 21; Length 25;

Best Local Similarity 100.0%; Pred. No. 5.3e-15;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGYGYQSVDRDPKPIVRSITLC 25

Db 1 CGYGYQSVDRDPKPIVRSITLC 25

RESULT 2

AY80018

ID AAY80018 standard; Peptide; 45 AA.

XX

AC AAY80018;

XX

DT 15-MAY-2000 (first entry)

DE IgE immunogenic peptide conjugate SEQ ID NO:25.

KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

OS Unidentified.

XX

PN WO9967293-A1.

XX

PD 29-DEC-1999.

XX

PF 21-JUN-1999; 99WO-US13959.

XX

PR 20-JUN-1998; 98US-0100287.

XX

PA (UNBI-) UNITED BIOMEDICAL INC.

XX

PI Wang CY, Walfield AM;

XX

DR WPI; 2000-160578/14.

XX

PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy

XX Claim 14; Page 76; 155pp; English.

XX

CC The present invention describes immunoglobulin E (IgE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils

CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

CC containing (I) are used for active immunisation against IgE-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell

CC target epitope, have increased immunogenicity and may include cyclic

CC constraints (disulfide bridge) to stabilise conformational features and

CC maximize cross-reactivity to the natural target. They induce safe

CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

CC acid sequences used in the exemplification of the present invention.

XX Sequence 45 AA;

5Q

Query Match 92.1%; Score 129; DB 21; Length 45;

Best Local Similarity 92.0%; Pred. No. 5.3e-13;

Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGYGYQSVDRDPKPIVRSITLC 25

Db 21 CGYGYQSVDRDPKPIVRSITLC 45

RESULT 3

AY80017

ID AAY80017 standard; Peptide; 46 AA.

XX

AC AAY80017;

XX

DT 15-MAY-2000 (first entry)

DE IgE immunogenic peptide conjugate SEQ ID NO:24.

KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

OS Unidentified.

XX

PN WO9967293-A1.

XX

PD 29-DEC-1999.

XX

PF 21-JUN-1999; 99WO-US13959.

XX

PR 20-JUN-1998; 98US-0100287.

XX

PA (UNBI-) UNITED BIOMEDICAL INC.

XX

PI Wang CY, Walfield AM;

XX

DR WPI; 2000-160578/14.

XX

PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

PT for immunization against allergy

XX Claim 14; Page 76; 155pp; English.

XX

CC The present invention describes immunoglobulin E (IgE)-CH3 domain

CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

CC and anti-asthmatic properties. (I) induces polyclonal antibodies

CC specific for a target effector site on the epsilon-heavy chain of IgE,

CC and so preventing triggering and activation of mast cells and basophils

CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,

CC containing (I) are used for active immunisation against IgE-mediated

CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

CC dermatitis. Nucleic acids that encode these compounds are useful for

CC recombinant production of corresponding peptides or in DNA vaccines.

CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.
XX
SQ Sequence 46 AA;

Query Match 92.1%; Score 129; DB 21; Length 46;
Best Local Similarity 92.0%; Pred. No. 5.4e-13;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGYGYQSIVDRPDPFKPIVRSITLC 25
||||| :|||||
DB 22 cgygyqsivdhpdpfkpivrsitkc 46

RESULT 4

AAAY17415
ID AAY17415 standard; Protein; 561 AA.
XX
AC AAY17415;
XX
DT 26-JUL-1999 (first entry)
XX
DE Mouse immunoglobulin E heavy chain.
XX
DE Mouse; immunoglobulin E; IgE; allergy; transgenic animal.
XX

Mus sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT /label= signal
FT Protein 20..561
FT /label= immunoglobulin_E_heavy_chain
XX
XX EP921189-A1.
XX
XX 09-JUN-1999.
XX
XX 13-NOV-1998; 98EP-0309340.
XX
XX 14-NOV-1997; 97JP-0313989.
XX

(SANYO) SANKYO CO LTD.
PA (TOKM-) TOKYO METROPOLITAN INST MEDICAL SCI.
XX
XX Karasuyama H, Matsuoaka K, Taya C, Yonemawa H;
XX
XX WPI; 1999-315404/27.
XX
XX N-PSDB; AAX61084.

Transgenic non-human animal allergy models

Claim 27; Page 24-27; 42pp; English.

XX The present invention describes a transgenic, non-human animal with its
CC genome altered to constitutively express a molecule having a constant
CC region which can bind an IgE receptor on mast cells in the animals, the
CC molecule having an immunoglobulin structure and being further capable of
CC specifically binding a predetermined antigen. The transgenic animal is
CC useful as a model for evaluating the activity and the ability of
CC substances i.e. with antiallergic activity, to affect any allergic
CC reaction caused in the animal by the administration of the antigen to
CC the animal, and applying the substance to be evaluated. The present
CC sequence is the mouse immunoglobulin E (IgE) heavy chain, given in
CC the present invention.
XX
SQ Sequence 561 AA;

Query Match 82.9%; Score 116; DB 20; Length 561;
Best Local Similarity 91.3%; Pred. No. 9.4e-10;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYGYQSIVDRPDPFKPIVRSITL 24
||||| :|||||
DB 419 ggygqvvdprdpfkpivrsitl 441

RESULT 5

AAAY79997
ID AAY79997 standard; Protein; 313 AA.
XX

AC AAY79997;
XX

DT 15-MAY-2000 (first entry)
XX

DE Mouse immunoglobulin E epsilon heavy chain SEQ ID NO:4.
XX

KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX

OS Mus sp.
XX

PN WO9967293-A1.
XX

PD 29-DEC-1999.
XX

PF 21-JUN-1999; 99WO-US13959.
XX

PR 20-JUN-1998; 98US-0100287.
XX

PA (UNBI-) UNITED BIOMEDICAL INC.
XX

PI Wang CY, Walfield AM;
XX

DR WPI; 2000-160578/14.
XX

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy -
XX

PS Example 1; Page 66-68; 155pp; English.
XX

CC The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.
XX
SQ Sequence 313 AA;

Query Match 80.7%; Score 113; DB 21; Length 313;
Best Local Similarity 95.5%; Pred. No. 1.4e-09;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYGYQSIVDRPDPFKPIVRSIT 23
||||| :|||||

Db 191 gygycivdrpdpfpkpvrsit 212

RESULT 6

AA80000
ID AAY80000 standard; Peptide; 25 AA.

XX AC AAY80000;

DT 15-MAY-2000 (first entry)

XX Optimised IgE-CH3 domain antigen peptide for rat IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

OS Rattus sp.

OS Synthetic.

PN WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

PA (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy -

PS Claim 1; Page 99; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

Query Match 75.7%; Score 106; DB 21; Length 25;

Best Local Similarity 80.0%; Pred. No. 1.1e-09;

Matches 20; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGYGYQSIIVDRPDPFPKPIVRSITLC 25

Db 1 cgegysrvdrhphfpkpvrsitkc 25

RESULT 7

AA79996
ID AAY79996 standard; Protein; 313 AA.

XX

AC AAY79996;

DT 15-MAY-2000 (first entry)

DE Rat Immunoglobulin E epsilon heavy chain SEQ ID NO:3.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

OS Rattus sp.

PN WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

PA (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy -

PS Example 1; Page 66-68; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.

XX Sequence 313 AA;

Query Match 60.7%; Score 85; DB 21; Length 313;

Best Local Similarity 77.3%; Pred. No. 3.3e-05;

Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GYGYQSIIVDRPDPFPKPIVRSIT 23

Db 190 gegycrvdrhphfpkpvrsit 211

RESULT 8

AA803643

ID AAB03643 standard; protein; 340 AA.

XX AAAB03643;

DT 22-NOV-2000 (first entry)

DE Rat IgE heavy chain constant regions 2, 3 and 4.

KW Rat; immunoglobulin E; IgE; vaccination; infection; allergy;
 XX asthma; eczema; immunogenic peptide.
 OS Rattus sp.

XX WO200025722-A2.

XX 11-MAY-2000.

XX 21-OCT-1999; 99WO-SE01896.

XX 02-NOV-1998; 98US-0106652.

XX 22-SEP-1999; 99US-0401636.

XX (RESI-) RESISTENTIA PHARM AB.

XX Hellman LT;

XX WPI; 2000-365342/31.

XX Immunogenic polypeptides useful for preventing the harmful effects of

XX immunoglobulin E in mammals -

XX Disclosure; Fig 1; 50pp; English.

XX The present sequence is an immunogenic peptide consisting of the
 CC heavy chain constant regions 2, 3 and 4' of the rat IgE. It was used to
 CC construct a number of immunogenic peptides which consisted of regions of
 CC IgE from different mammals, which appear to cause a stronger polyclonal
 CC anti-self IgE response than peptides consisting of the same regions from
 CC one mammal. Immunogenic peptides, particularly those consisting of
 CC different heavy chain constant regions, can be used for vaccination
 CC in humans, against bacterial and viral infections and allergies, such
 CC as asthma, fur, pollen and food allergies and eczema.

XX Sequence 340 AA;

Query Match 60.7%; Score 85; DB 21; Length 340;
 Best Local Similarity 77.3%; Pred. No. 3.6e-05;

Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GYGQSVDRDPFPKPIVRSIT 23

Db 195 gegyqcrvdhphfhpivrsit 216

RESULT 9

AAB06206

ID AAB06206 standard; protein; 341 AA.

XX AAB06206;

XX 22-NOV-2000 (first entry)

XX Immunogenic peptide consisting of opossum CH2, rat CH3 and opossum CH4.

XX Rat; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;

XX asthma; eczema; immunogenic peptide.

XX Chimeric - Didelphis virginiana.

XX Chimeric - Rattus sp.

XX WO200025722-A2.

XX 11-MAY-2000.

XX 21-OCT-1999; 99WO-SE01896.

XX 02-NOV-1998; 98US-0106652.

XX 22-SEP-1999; 99US-0401636.

XX (RESI-) RESISTENTIA PHARM AB.

XX Hellman LT;

XX WPI; 2000-365342/31.

XX Immunogenic polypeptides useful for preventing the harmful effects of

XX immunoglobulin E in mammals -

XX Disclosure; Fig 2; 50pp; English.

XX The present sequence is an immunogenic peptide consisting of the heavy
 CC chain constant regions 2 and 4 of the opossum IgE and the heavy chain
 CC constant region 3 from the rat. It was shown to cause a stronger
 CC polyclonal anti-self IgE response than peptides consisting of the same
 CC regions from one mammal. Immunogenic peptides, particularly those
 CC consisting of different heavy chain constant regions, can be used for
 CC vaccination in humans, against bacterial and viral infections and
 CC allergies, such as asthma, fur, pollen and food allergies and eczema.

XX Sequence 341 AA;

Query Match 60.7%; Score 85; DB 21; Length 341;

Best Local Similarity 77.3%; Pred. No. 3.6e-05;

Matches 17; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GYGQSVDRDPFPKPIVRSIT 23

Db 206 gegyqcrvdhphfhpivrsit 227

RESULT 10

AAy80077

ID AAY80077 standard; Peptide; 25 AA.

XX AC AAY80077;

XX DT 15-MAY-2000 (first entry)

XX Optimised IgE-CH3 domain antigen peptide for horse IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Equus caballus.

XX Synthetic.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -

XX Claim 1; Page 146; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils

CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;

Query Match 52.9%; Score 74; DB 21; Length 25;
 Best Local Similarity 52.0%; Pred. No. 0.0001;
 Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 CGYGQSIIVDRDPFKPIVRSITLC 25
 || | | | | | | | | | | | | | | | |
 Db 1 cgetyktvshpdlprevvrslakc 25

RESULT 11

AAY79999
 ID AAY79999 standard; Peptide; 25 AA.

XX
 AC AAY79999;

XX
 DT 15-MAY-2000 (first entry)

XX
 DE Optimised IgE-CH3 domain antigen peptide for dog IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX
 OS Canis sp.
 OS Synthetic.

XX
 PN WO9967293-A1.

XX
 PD 29-DEC-1999.

XX
 PF 21-JUN-1999; 99WO-US13959.

XX
 PR 20-JUN-1998; 98US-0100287.

XX
 PA (UNBI-) UNITED BIOMEDICAL INC.

XX
 PI Wang CY, Walfield AM;

XX
 DR WPI; 2000-160578/14.

XX
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -

XX
 PS Claim 1; Page 99; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope

CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 25 AA;

Query Match 47.1%; Score 66; DB 21; Length 25;
 Best Local Similarity 56.0%; Pred. No. 0.0018;
 Matches 14; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CGYGQSIIVDRDPFKPIVRSITLC 25
 || | | | | | | | | | | | | | | | |
 Db 1 cgetyysrvthphlpkdivrslakc 25

RESULT 12

AAY80019
 ID AAY80019 standard; Peptide; 45 AA.

XX
 AC AAY80019;

XX
 DT 15-MAY-2000 (first entry)

XX
 DE IgE immunogenic peptide conjugate SEQ ID NO:26.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX
 OS Unidentified.

XX
 PN WO9967293-A1.

XX
 PD 29-DEC-1999.

XX
 PF 21-JUN-1999; 99WO-US13959.

XX
 PR 20-JUN-1998; 98US-0100287.

XX
 PA (UNBI-) UNITED BIOMEDICAL INC.

XX
 PI Wang CY, Walfield AM;

XX
 DR WPI; 2000-160578/14.

XX
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -

XX
 PS Claim 14; Page 76; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.

SQ Sequence 45 AA;

Query Match 47.1%; Score 66; DB 21; Length 45;
Best Local Similarity 56.0%; Pred. No. 0.0035;
Matches 14; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CGYGYQSIVDRDPFKPIVRSITLC 25
II III I IIIII I
DB 21 cgetysrvthphlpkdivrsiack 45

RESULT 13

AAV80083
ID AAY80083 standard; Peptide: 45 AA.

XX AC AAY80083;

DT 15-MAY-2000 (first entry)

DE IgE immunogenic peptide conjugate SEQ ID NO:90.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

PN WO9967293-A1.

PD 29-DEC-1999.

PF 21-JUN-1999; 99WO-US13959.

PR 20-JUN-1998; 98US-0100287.

PA (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY, Walfield AM;

DR WPI; 2000-160578/14.

PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
for immunization against allergy

PS Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
and anti-asthmatic properties. (I) induces polyclonal antibodies
specific for a target effector site on the epsilon-heavy chain of IgE,
and so preventing triggering and activation of mast cells and basophils
and downregulation of IgE synthesis. Conjugates, or fusion peptides,
containing (I) are used for active immunisation against IgE-mediated
allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
dermatitis. Nucleic acids that encode these compounds are useful for
recombinant production of corresponding peptides or in DNA vaccines.
Conjugates of (I) that include a promiscuous T helper cell epitope
(functional in genetically diverse subjects), in addition to a B cell
target epitope, have increased immunogenicity and may include cyclic
constraints (disulfide bridge) to stabilise conformational features and
maximize cross-reactivity to the natural target. They induce safe
(non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
acid sequences used in the exemplification of the present invention.

SQ Sequence 45 AA;

Query Match 47.1%; Score 66; DB 21; Length 45;
Best Local Similarity 56.0%; Pred. No. 0.0035;
Matches 14; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CGYGYQSIVDRDPFKPIVRSITLC 25
II III I IIIII I
DB 21 cgetysrvthphlpkdivrsiack 45

RESULT 14

AAV80020
ID AAY80020 standard; Peptide: 46 AA.

XX AC AAY80020;

DT 15-MAY-2000 (first entry)

DE IgE immunogenic peptide conjugate SEQ ID NO:27.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

PN WO9967293-A1.

PD 29-DEC-1999.

PF 21-JUN-1999; 99WO-US13959.

PR 20-JUN-1998; 98US-0100287.

PA (UNBI-) UNITED BIOMEDICAL INC.

PI Wang CY, Walfield AM;

DR WPI; 2000-160578/14.

PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
for immunization against allergy

PS Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
and anti-asthmatic properties. (I) induces polyclonal antibodies
specific for a target effector site on the epsilon-heavy chain of IgE,
and so preventing triggering and activation of mast cells and basophils
and downregulation of IgE synthesis. Conjugates, or fusion peptides,
containing (I) are used for active immunisation against IgE-mediated
allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
dermatitis. Nucleic acids that encode these compounds are useful for
recombinant production of corresponding peptides or in DNA vaccines.
Conjugates of (I) that include a promiscuous T helper cell epitope
(functional in genetically diverse subjects), in addition to a B cell
target epitope, have increased immunogenicity and may include cyclic
constraints (disulfide bridge) to stabilise conformational features and
maximize cross-reactivity to the natural target. They induce safe
(non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
acid sequences used in the exemplification of the present invention.

SQ Sequence 46 AA;

Query Match 47.1%; Score 66; DB 21; Length 46;
Best Local Similarity 56.0%; Pred. No. 0.0035;
Matches 14; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 CGYGYQSIVDRDPFKPIVRSITLC 25
II III I IIIII I
DB 22 cgetysrvthphlpkdivrsiack 46

RESULT 15

AAV80081

ID AAY80081 standard; Peptide; 57 AA.
 XX
 AC AAY80081;
 XX
 DT 15-MAY-2000 (first entry)
 XX
 DE IgE immunogenic peptide conjugate SEQ ID NO:88.
 XX
 KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX
 OS Unidentified.
 XX
 PN WO9967293-A1.
 XX
 PD 29-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-US13959.
 XX
 PR 20-JUN-1998; 98US-0100287.
 XX
 PA (UNBI-) UNITED BIOMEDICAL INC.
 XX
 PI Wang CY, Walfield AM;
 XX
 DR WPI: 2000-160578/14.
 XX
 PT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -
 XX
 PS Claim 14; Page 77; 155pp; English.
 XX
 CC The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC Conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 57 AA;

Query Match 47.1%; Score 66; DB 21; Length 57;
 Best Local Similarity 56.0%; Pred. No. 0.0045;
 Matches 14; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
 QY 1 CGYGYQSIYDRDPFKPIVRSITLC 25
 DB 33 cgetyysrvthphlpkdivrsiakc 57

Search completed: March 4, 2002, 12:55:50
 Job time: 297 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model.

Run on: March 4, 2002, 12:55:50 ; Search time 82.62 Seconds
(without alignments)
22.414 Million cell updates/sec

Title: US-09-701-623C-84
Perfect score: 135
Sequence: 1 CGETYKSTVSHPDLPREVRSIAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101:*

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 135 | 100.0 | 25 | 21 | AA1980077 |
| 2 | 105 | 77.8 | 25 | 21 | AA1979999 |
| 3 | 105 | 77.8 | 45 | 21 | AA1980019 |
| 4 | 105 | 77.8 | 45 | 21 | AA1980083 |
| 5 | 105 | 77.8 | 46 | 21 | AA1980020 |
| 6 | 105 | 77.8 | 57 | 21 | AA1980081 |
| 7 | 105 | 77.8 | 62 | 21 | AA1980080 |
| 8 | 105 | 77.8 | 63 | 21 | AA1980084 |
| 9 | 94 | 69.6 | 25 | 21 | AA1979998 |
| 10 | 94 | 69.6 | 25 | 21 | AA191212 |
| 11 | 94 | 69.6 | 25 | 21 | AA1986602 |

| | | | | | | |
|----|----|------|-----|----|-----------|--------------------|
| 12 | 94 | 69.6 | 42 | 21 | AA1980014 | IgE immunogenic pe |
| 13 | 94 | 69.6 | 42 | 21 | AA191215 | Modified MvF Th ep |
| 14 | 94 | 69.6 | 42 | 21 | AA191216 | Modified MvF Th ep |
| 15 | 94 | 69.6 | 42 | 21 | AA191217 | Modified MvF Th ep |
| 16 | 94 | 69.6 | 42 | 21 | AA1986604 | Peptide sequence o |
| 17 | 94 | 69.6 | 45 | 21 | AA1980007 | IgE-CH3 domain ant |
| 18 | 94 | 69.6 | 45 | 21 | AA191218 | Modified HBV surfa |
| 19 | 94 | 69.6 | 45 | 21 | AA1986605 | Peptide sequence o |
| 20 | 94 | 69.6 | 46 | 21 | AA1980011 | IgE immunogenic pe |
| 21 | 94 | 69.6 | 46 | 21 | AA191213 | Modified MvF Th ep |
| 22 | 94 | 69.6 | 46 | 21 | AA191214 | Modified MvF Th ep |
| 23 | 94 | 69.6 | 46 | 21 | AA1986603 | Peptide sequence o |
| 24 | 94 | 69.6 | 56 | 21 | AA1980016 | IgE immunogenic pe |
| 25 | 94 | 69.6 | 59 | 21 | AA1980010 | IgE immunogenic pe |
| 26 | 94 | 69.6 | 60 | 21 | AA1980013 | IgE immunogenic pe |
| 27 | 94 | 69.6 | 60 | 21 | AA1980015 | IgE immunogenic pe |
| 28 | 94 | 69.6 | 63 | 21 | AA1980008 | IgE-CH3 domain ant |
| 29 | 94 | 69.6 | 63 | 21 | AA1980012 | IgE immunogenic pe |
| 30 | 94 | 69.6 | 63 | 21 | AA191219 | Inv epitope/modifi |
| 31 | 94 | 69.6 | 63 | 21 | AA1986606 | Peptide sequence o |
| 32 | 89 | 65.9 | 25 | 21 | AA1980000 | Optimised IgE-CH3 |
| 33 | 89 | 65.9 | 45 | 21 | AA1980018 | IgE immunogenic pe |
| 34 | 89 | 65.9 | 46 | 21 | AA1980017 | IgE immunogenic pe |
| 35 | 87 | 64.4 | 60 | 21 | AA1980078 | IgE immunogenic pe |
| 36 | 82 | 60.7 | 124 | 18 | AA1924097 | Partial canine imm |
| 37 | 82 | 60.7 | 312 | 21 | AA1979995 | Dog immunoglobulin |
| 38 | 82 | 60.7 | 341 | 21 | AA1986208 | Immunogenic peptid |
| 39 | 82 | 60.7 | 417 | 18 | AA1923067 | Canine IgE heavy c |
| 40 | 82 | 60.7 | 426 | 17 | AA1977553 | Canine IgE. Canis |
| 41 | 81 | 60.0 | 345 | 21 | AA1986207 | Immunogenic peptid |
| 42 | 74 | 54.8 | 25 | 21 | AA1980001 | Optimised IgE-CH3 |
| 43 | 72 | 53.3 | 110 | 14 | AA1933324 | Variant IgE - muta |
| 44 | 71 | 52.6 | 106 | 20 | AA1942620 | Human IgE Fcpsiilo |
| 45 | 71 | 52.6 | 110 | 14 | AA1933304 | IgE Fc epsilon 3. |

ALIGNMENTS

RESULT 1

AA1980077
ID AA1980077 standard; Peptide; 25 AA.

XX AC AA1980077;

XX AC
DT 15-MAY-2000 (first entry)

XX DE Optimised IgE-CH3 domain antigen peptide for horse IgE.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Equus caballus.

XX OS Synthetic

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PI Wang CY, Walfield AM;

XX DR WPI; 2000-160578/14.

XX PT New antigenic peptide from the CH3 domain of Immunoglobulin E, fusions
PT for immunization against allergy

XX PS Claim 1; Page 146; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain

XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies

XX CC specific for a target effector site on the epsilon-heavy chain of IgE,

XX CC and so preventing triggering and activation of mast cells and basophils

XX CC containing (I) are used for active immunisation against IgE-mediated

XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

XX CC dermatitis. Nucleic acids that encode these compounds are useful for

XX CC recombinant production of corresponding peptides or in DNA vaccines.

XX CC Conjugates of (I) that include a promiscuous T helper cell epitope

XX CC (functional in genetically diverse subjects), in addition to a B cell

XX CC target epitope, have increased immunogenicity and may include cyclic

XX CC constraints (disulfide bridge) to stabilise conformational features and

XX CC maximize cross-reactivity to the natural target. They induce safe

XX CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

XX CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 25 AA;

Query Match 100.0%; Score 135; DB 21; Length 25;

Best Local Similarity 100.0%; Pred. No. 1.2e-14;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHDLPREVVRSIAC 25

Db 1 cgetykstvshdiprevvrslac 25

||||| | | | | | | | | | | | | | | | | | | | |

RESULT 2

AAY79999

ID AAY79999 standard; Peptide; 25 AA.

XX AC AAY79999;

XX DT 15-MAY-2000 (first entry)

XX DE Optimised IgE-CH3 domain antigen peptide for dog IgE.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;

XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Canis sp.

XX OS Synthetic.

XX PN WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PT Wang CY, Walfield AM;

XX PI WPI; 2000-160578/14.

XX DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

XX PT for immunization against allergy -

XX PS Claim 1; Page 99; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain

XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies

XX CC specific for a target effector site on the epsilon-heavy chain of IgE,

XX CC and so preventing triggering and activation of mast cells and basophils

XX CC containing (I) are used for active immunisation against IgE-mediated

XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

XX CC dermatitis. Nucleic acids that encode these compounds are useful for

XX CC recombinant production of corresponding peptides or in DNA vaccines.

XX CC Conjugates of (I) that include a promiscuous T helper cell epitope

XX CC (functional in genetically diverse subjects), in addition to a B cell

XX CC target epitope, have increased immunogenicity and may include cyclic

XX CC constraints (disulfide bridge) to stabilise conformational features and

XX CC maximize cross-reactivity to the natural target. They induce safe

XX CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino

XX CC acid sequences used in the exemplification of the present invention.

XX SQ Sequence 25 AA;

Query Match 77.8%; Score 105; DB 21; Length 25;

Best Local Similarity 72.0%; Pred. No. 7.3e-10;

Matches 18; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHDLPREVVRSIAC 25

Db 1 cgetykstvshdiprevvrslac 25

||||| | | | | | | | | | | | | | | | | | | | |

RESULT 3

AAY80019

ID AAY80019 standard; Peptide; 45 AA.

XX AC AAY80019;

XX DT 15-MAY-2000 (first entry)

XX DE IgE immunogenic peptide conjugate SEQ ID NO:26.

XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

XX KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

XX KW antibody; allergy; allergic disease; immunisation; anti-allergic;

XX KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX OS Unidentified.

XX OS WO9967293-A1.

XX PD 29-DEC-1999.

XX PF 21-JUN-1999; 99WO-US13959.

XX PR 20-JUN-1998; 98US-0100287.

XX PA (UNBI-) UNITED BIOMEDICAL INC.

XX PT Wang CY, Walfield AM;

XX PI WPI; 2000-160578/14.

XX DR New antigenic peptide from the CH3 domain of immunoglobulin E, fusions

XX PT for immunization against allergy -

XX PS Claim 14; Page 76; 155pp; English.

XX CC The present invention describes immunoglobulin E (IgE)-CH3 domain

XX CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic

XX CC and anti-asthmatic properties. (I) induces polyclonal antibodies

XX CC specific for a target effector site on the epsilon-heavy chain of IgE,

XX CC and so preventing triggering and activation of mast cells and basophils

XX CC containing (I) are used for active immunisation against IgE-mediated

XX CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy

XX CC dermatitis. Nucleic acids that encode these compounds are useful for

XX CC recombinant production of corresponding peptides or in DNA vaccines.

| | | | | |
|--------------------------|--------|--------------------|--------|------------|
| Query Match | 77.8%; | Score 105; | DB 21; | Length 45; |
| Best Local Similarity | 72.0%; | Pred. No. 1.5e-09; | | |
| Matches 18; Conservative | 4; | Mismatches 3; | Indels | |

```
Qy 1 CGETVKSTVSHPDLPREVRSIAKC 25
    ||||| + | : | | : : |||||
Db 22 cqetyvsrv:hphlpkdivrsiac 46
```

| | |
|----------|-------------------------------------------------|
| RESULT | 6 |
| RAY80081 | |
| ID | RAY80081 standard; Peptide; 57 AA. |
| XX | |
| XX | |
| AC | RAY80081; |
| DT | |
| DT | 15-MAY-2000 (first entry) |
| XX | |
| XX | |
| DE | |
| DE | IgE immunogenic peptide conjugate SEQ ID NO:88. |

RESULT 7

DE IgE immunogenic peptide conjugate SEQ ID NO:91.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW antibody; allergy; allergic disease; immunisation; anti-allergic;
KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX Unidentified.

XX WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy

XX Claim 14; Page 77; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.

XX Sequence 63 AA;

Query Match 77.8%; Score 105; DB 21; Length 63;

Best Local Similarity 72.0%; Pred. No. 2.2e-09;

Matches 18; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHPDLPREVRSIAKC 25

Db 39 cgetysrvthpdkdivrsiack 63

RESULT 9

AA79998
ID AAY79998 standard; Peptide; 25 AA.

XX AAY79998;

XX 15-MAY-2000 (first entry)

DE Optimised IgE-CH3 domain antigen peptide for human IgE.

XX Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;

KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;

KW antibody; allergy; allergic disease; immunisation; anti-allergic;

KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.

XX

OS Homo sapiens.

OS Synthetic.

PN WO9967293-A1.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13959.

XX 20-JUN-1998; 98US-0100287.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY, Walfield AM;

XX WPI; 2000-160578/14.

XX New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT for immunization against allergy

XX Claim 1; Page 21; 155pp; English.

XX The present invention describes immunoglobulin E (IgE)-CH3 domain
CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
CC and anti-asthmatic properties. (I) induces polyclonal antibodies
CC specific for a target effector site on the epsilon-heavy chain of IgE,
CC and so preventing triggering and activation of mast cells and basophils
CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC containing (I) are used for active immunisation against IgE-mediated
CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC dermatitis. Nucleic acids that encode these compounds are useful for
CC recombinant production of corresponding peptides or in DNA vaccines.
CC Conjugates of (I) that include a promiscuous T helper cell epitope
CC (functional in genetically diverse subjects), in addition to a B cell
CC target epitope, have increased immunogenicity and may include cyclic
CC constraints (disulfide bridge) to stabilise conformational features and
CC maximize cross-reactivity to the natural target. They induce safe
CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC acid sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

Query Match 69.6%; Score 94; DB 21; Length 25;

Best Local Similarity 64.0%; Pred. No. 4.1e-08;

Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHPDLPREVRSIAKC 25

Db 1 cgetysrvthpdkdivrsiack 25

RESULT 10

AA791212

ID AAY91212 standard; peptide; 25 AA.

XX AAY91212;

XX 22-MAY-2000 (first entry)

DE Modified human IgE CH3 domain, SEQ ID NO:92.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;

KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;

KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;

KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;

KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;

KW Plasmodium falciparum; circumsporozoite; antimalarial; CPTP;

KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX Homo sapiens.

OS Synthetic.

XX

PN WO9966957-A2.
PD 29-DEC-1999.
XX 21-JUN-1999; 99WO-US13975.
XX 20-JUN-1998; 98US-0100412.
XX (UNBI-) UNITED BIOMEDICAL INC.
XX Wang CY;
XX WPI: 2000-160564/14.
XX New artificial T helper cell epitope and derived immunogens with target
XX antigenic site, for immunization against e.g. malaria, arteriosclerosis
XX or human immune deficiency virus -
XX
XX Example 6; Page 40; 129pp; English.
XX
XX The invention relates to novel promiscuous T helper cell epitopes (Th),
XX and immunogenic peptides comprising the Th epitopes of the invention
XX along with B cell epitopes. The Th epitopes and peptide immunogens
XX containing them, are used to induce a T helper cell response,
XX specifically against Plasmodium falciparum, cholesterol ester transport
XX protein (CEPT) or HIV epitopes, but more generally against any pathogen,
XX immunoreactive self-antigen or tumour antigen. The Th epitopes and
XX peptide immunogens may be used for prevention and/or treatment of
XX infections (HIV, foot-and-mouth disease or malaria); for cancer
XX immunotherapy; for inhibition of the action of luteinising hormone
XX releasing hormone (LHRH) for contraception, treatment of hormone-
XX dependent cancer, prevention of boar taint in meat, and
XX immunocastration); for promoting the growth of animals; or for
XX treating allergies or arteriosclerosis. Incorporation of a promiscuous
XX Th (functional in genetically diverse subjects) into an immunogen
XX improves capacity to induce a strong T helper cell-mediated immune
XX response, resulting in production of antibodies against a target
XX antigen. Th can replace carrier proteins and pathogen-derived T helper
XX epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
XX from the measles virus F (MV) protein and sequences AAY91122-Y91142,
XX AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
XX MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
XX from hepatitis B virus (HBV) surface antigen, and sequences
XX AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
XX AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
XX comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
XX is the LHRH target antigenic peptide used in these LHRH antigenic
XX peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
XX peptides comprising somatostatin and a Th epitope. Somatostatin
XX immunogens may be used to promote growth in livestock. AAY91208 is a
XX human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
XX epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
XX infection of T cells. AAY90212 is a modified version of a human IgE
XX (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
XX antigenic peptides which may be used in the treatment of allergies.
XX AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
XX VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
XX epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
XX antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
XX epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
XX CPTP-derived peptides and AAY91232-Y91241 are immunogens comprising a
XX CPTP peptide and a Th epitope which may be used to prevent or treat
XX arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
XX are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
XX AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
XX epitope which may be used as a component in an anti-HIV-1 vaccine.
XX AAY91198 and AAY91199 are respectively an immunostimulatory invasive
XX protein epitope from *Yersinia* species, and hinge spacer peptide, both of
XX which may optionally be used in the antigenic peptides of the
XX invention.
XX
XX Sequence 25 AA:
SQ

Query Match 69.6%; Score 94; DB 21; Length 25;
Best Local Similarity 64.0%; Pred. No. 4.1e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
OY 1 CGETYKSTVSHPDLPREVVRISAKC 25
DB 1 cgetyqsrvtphlpralmrsttkc 25
IIIIII:1:111111:111111

RESULT 11
AAY68602
ID AAY68602 standard; peptide; 25 AA.
XX
XX AC AAY68602;
XX DT 05-MAY-2000 (first entry)
XX DE Peptide sequence of the invention.
XX
XX KW Helper T cell epitope; peptide immunogen; LHRH;
XX luteinising hormone-releasing hormone; spermatogenesis; ovulation;
XX oestrus; sexual development; sex hormone; promiscuous T helper epitope;
XX vaccine; contraceptive; hormone-dependent tumour; prostate cancer;
XX breast cancer; endometriosis; boar taint; meat quality; immunocastration.
XX
XX OS Unidentified.
XX PN WO9966952-A1.
XX PD 29-DEC-1999.
XX PF 21-JUN-1999; 99WO-US13960.
XX PR 20-JUN-1998; 98US-0100414.
XX PA (UNBI-) UNITED BIOMEDICAL INC.
XX PI Wang CY;
XX WPI: 2000-160562/14.
XX
XX New peptide immunogen containing luteinising hormone-releasing hormone
XX antigen site and helper T cell epitope, for e.g. contraception and
XX treatment of cancer
XX
XX PS Disclosure; Page 92; 102pp; English.
XX
XX The specification describes peptide immunogens comprising a
XX synthetic helper T cell (Th) epitope and a target antigen, luteinising
XX hormone-releasing hormone (LHRH). The peptide immunogens cause
XX induction of a specific immune response to LHRH which is involved in
XX regulation of spermatogenesis, ovulation, oestrus, sexual development
XX and secretion of sex hormones. Provision of a promiscuous T helper
XX epitope (which is functional in genetically diverse subjects) provides
XX optimum immunogenicity to the B cell epitopes of the target antigen and
XX thus high antibody titres against the target antigen. The peptide
XX immunogens of the invention are used to vaccinate against mammalian LHRH,
XX for use as (reversible) contraceptive; control of hormone-dependent
XX tumours (cancer of prostate or breast, also endometriosis); to prevent
XX boar taint (and improve meat quality) and for immunocastration. The
XX present sequence appears in the specification.
XX
XX SQ Sequence 25 AA;
SQ

Query Match 69.6%; Score 94; DB 21; Length 25;
Best Local Similarity 64.0%; Pred. No. 4.1e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
OY 1 CGETYKSTVSHPDLPREVVRISAKC 25
DB 1 cgetyqsrvtphlpralmrsttkc 25
IIIIII:1:111111:111111

RESULT 12
 AAY80014
 ID AAY80014 standard; Peptide; 42 AA.
 XX AC AAY80014;
 XX DT 15-MAY-2000 (first entry)
 XX DE IgE immunogenic peptide conjugate SEQ ID NO:21.
 XX KW Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
 KW immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
 KW antibody; allergy; allergic disease; immunisation; anti-allergic;
 KW anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
 XX OS Unidentified.
 XX PN WO9967293-A1.
 XX PD 29-DEC-1999.
 XX PF 21-JUN-1999; 99WO-US13959.
 XX PR 20-JUN-1998; 98US-0100287.
 XX PA (UNBI-) UNITED BIOMEDICAL INC.
 XX PI Wang CY, Walfield AM;
 XX WPI; 2000-160578/14.
 XX DT New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
 PT for immunization against allergy -
 PS Claim 14; Page 76; 155pp; English.
 XX The present invention describes immunoglobulin E (IgE)-CH3 domain
 CC antigenic peptides (I). (I) have anti-allergic, anti-anaphylactic
 CC and anti-asthmatic properties. (I) induces polyclonal antibodies
 CC specific for a target effector site on the epsilon-heavy chain of IgE,
 CC and so preventing triggering and activation of mast cells and basophils
 CC and downregulation of IgE synthesis. Conjugates, or fusion peptides,
 CC containing (I) are used for active immunisation against IgE-mediated
 CC allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
 CC dermatitis. Nucleic acids that encode these compounds are useful for
 CC recombinant production of corresponding peptides or in DNA vaccines.
 CC conjugates of (I) that include a promiscuous T helper cell epitope
 CC (functional in genetically diverse subjects), in addition to a B cell
 CC target epitope, have increased immunogenicity and may include cyclic
 CC constraints (disulfide bridge) to stabilise conformational features and
 CC maximize cross-reactivity to the natural target. They induce safe
 CC (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
 CC acid sequences used in the exemplification of the present invention.
 XX SQ Sequence 42 AA:
 Query Match 69.6%; Score 94; DB 21; Length 42;
 Best Local Similarity 64.0%; Pred. No. 7.7e-08;
 Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 QY 1 CGETYKSTVSHPDLPREVRSIAKC 25
 Db 18 cgetyqsrvtphpralmrsttkc 42
 RESULT 13
 AAY91215
 ID AAY91215 standard; peptide; 42 AA.
 XX AC AAY91215;

XX 22-MAY-2000 (first entry)
 XX Modified MVF Th epitope/IgE CH3 domain, SEQ ID NO:95.
 XX Promiscuous T-cell epitope; measles virus F protein; MVF;
 KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
 KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
 KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
 KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
 KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
 KW cholesteryl ester transport protein; anti-arteriosclerotic.
 XX Chimeric - Measles virus.
 OS Chimeric - Homo sapiens.
 XX WO9966957-A2.
 XX PD 29-DEC-1999.
 XX PF 21-JUN-1999; 99WO-US13975.
 XX PR 20-JUN-1998; 98US-0100412.
 XX PA (UNBI-) UNITED BIOMEDICAL INC.
 XX PI Wang CY;
 XX WPI; 2000-160564/14.
 XX New artificial T helper cell epitope and derived immunogens with target
 PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
 PT or human immune deficiency virus -
 XX Example 6; Page 98-99; 129pp; English.
 XX The invention relates to novel promiscuous T helper cell epitopes (Th),
 CC and immunogenic peptides comprising the Th epitopes of the invention
 CC along with B cell epitopes. The Th epitopes and peptide immunogens
 CC containing them, are used to induce a T helper cell response,
 CC specifically against Plasmodium falciparum, cholesteryl ester transport
 CC protein (CERP) or HIV epitopes, but more generally against any pathogen,
 CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
 CC peptide immunogens may be used for prevention and/or treatment of
 CC infections (HIV, foot-and-mouth disease or malaria); for cancer
 CC immunotherapy; for inhibition of the action of luteinising hormone
 CC releasing hormone (LHRH) for contraception, treatment of hormone-
 CC dependent cancer, prevention of boar taint in meat, and
 CC immunocastration; for promoting the growth of animals; or for
 CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
 CC Th (functional in genetically diverse subjects) into an immunogen
 CC improves capacity to induce a strong T helper cell-mediated immune
 CC response, resulting in production of antibodies against a target
 CC antigen. Th can replace carrier proteins and pathogen-derived T helper
 CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
 CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
 CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
 CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
 CC from hepatitis B virus (HBV) surface antigen, and sequences
 CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
 CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
 CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
 CC is the LHRH target antigenic peptide used in these LHRH antigenic
 CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
 CC peptides comprising somatostatin and a Th epitope. Somatostatin
 CC immunogens may be used to promote growth in livestock. AAY91208 is a
 CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
 CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
 CC infection of T cells. AAY90212 is a modified version of a human IgE
 CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
 CC antigenic peptides which may be used in the treatment of allergies.
 CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
 CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th

epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a CERP peptide and a Th epitope which may be used to prevent or treat arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257 are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as a component in an anti-HIV-1 vaccine. AAY91198 and AAY91199 are respectively an immunostimulatory invasive protein epitope from Yersinia species, and hinge spacer peptide, both of which may optionally be used in the antigenic peptides of the invention.

Sequence 42 AA;

Query Match 69.6%; Score 94; DB 21; Length 42;
Best Local Similarity 64.0%; Pred. No. 7.7e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHPDLPREVRSIAKC 25

Db 18 cgetyqsrvtphpralmrsttkc 42

RESULT 14

AAY91216

ID AAY91216 standard; peptide: 42 AA.

XX AC AAY91216;

XX 22-MAY-2000 (first entry)

DE Modified MVF Th epitope/IgE CH3 domain, SEQ ID NO:96.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX Chimeric - Measles virus.

OS Chimeric - Homo sapiens.

XX WO9966957-A2.

XX 29-DEC-1995.

XX 21-JUN-1999; 99WO-US13975.

XX 20-JUN-1998; 98US-0100412.

XX (UNBI-) UNITED BIOMEDICAL INC.

PA Wang CY;

PI WPT: 2000-160564/14.

XX New artificial T helper cell epitope and derived immunogens with target
XX antigenic site, for immunization against e.g. malaria, arteriosclerosis
XX or human immune deficiency virus

PS Example 6; Page 99; 139pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CERP) or HIV epitopes, but more generally against any pathogen,

CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration; for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a
CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y91211 are MVH Th
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AAY90212 is a modified version of a human IgE
CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
CC epitope and may be used in a malaria vaccine. AAY91228-Y91231 represent
CC CERP-derived peptides and AAY91232-Y91241 are immunogens comprising a
CC CERP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AAY91198 and AAY91199 are respectively an immunostimulatory invasive
CC protein epitope from Yersinia species, and hinge spacer peptide, both of
CC which may optionally be used in the antigenic peptides of the
CC invention.

XX Sequence 42 AA;

Query Match 69.6%; Score 94; DB 21; Length 42;

Best Local Similarity 64.0%; Pred. No. 7.7e-08;

Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHPDLPREVRSIAKC 25

Db 18 cgetyqsrvtphpralmrsttkc 42

RESULT 15

AAY91217

ID AAY91217 standard; peptide: 42 AA.

XX AC AAY91217;

XX 22-MAY-2000 (first entry)

DE Modified MVF Th epitope/IgE CH3 domain, SEQ ID NO:97.

XX Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CERP;

KW cholesteryl ester transport protein; anti-arteriosclerotic.

XX Chimeric - Measles virus.
OS Chimeric - Homo sapiens.

XX WO9966957-A2.

XX 29-DEC-1999.

XX 21-JUN-1999; 99WO-US13975.

XX 20-JUN-1998; 98US-0100412.

XX (UNBI-) UNITED BIOMEDICAL INC.

XX Wang CY;

XX WPI; 2000-160564/14.

XX New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus

XX Example 6; Page 99; 129pp; English.

XX The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone-
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration; for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target
CC antigen. Th can replace carrier proteins and pathogen-derived T helper
CC epitopes. Sequence AAY91121 represents a promiscuous T helper epitope
CC from the measles virus F (MVF) protein and sequences AAY91122-Y91142,
CC AAY91226 and AAY91245-Y91246 represent synthetic Th epitopes based on the
CC MVF Th epitope. Sequence AAY91143 represents a promiscuous Th epitope
CC from hepatitis B virus (HBV) surface antigen, and sequences
CC AAY91144-Y91155 are synthetic epitopes derived from this HBV epitope.
CC AAY91156-Y91196, AAY91227 and AAY91242-Y91244 are antigenic peptides
CC comprising an LHRH sequence joined to a promiscuous Th epitope. AAY91197
CC is the LHRH target antigenic peptide used in these LHRH antigenic
CC peptides. AAY91200 is somatostatin, and AAY91201-Y91207 are antigenic
CC peptides comprising somatostatin and a Th epitope. Somatostatin
CC immunogens may be used to promote growth in livestock. AAY91208 is a
CC human CD4 CDR2-like domain antigenic site, and AAY91209-Y90211 are MVH Th
CC epitope/CD4 CDR2 antigenic peptides which may be used to prevent HIV
CC infection of T cells. AAY90212 is a modified version of a human IGE
CC (immunoglobulin E) CH3 domain, and AAY90213-Y90219 are Th epitope/IgE CH3
CC antigenic peptides which may be used in the treatment of allergies.
CC AAY91220 is a peptide derived from foot and mouth disease virus (FMDV)
CC VP1 capsid protein and AAY91221-Y91222 comprise this peptide and a Th
CC epitope. AAY91223 is a Plasmodium falciparum circumsporozoite (CS) target
CC antigen, and AAY91224-Y91225 comprise the CS antigen and an MVF Th
CC epitope, and may be used in a malaria vaccine. AAY91228-Y91231 represent
CC CETP-derived peptides and AAY91232-Y91241 are immunogens comprising a
CC CETP peptide and a Th epitope which may be used to prevent or treat
CC arteriosclerosis and cardiovascular disease. AAY91247 and AAY91252-Y91257
CC are HIV-1 neutralising B-cell epitopes, and AAY91248-Y91251 and
CC AAY91258-Y91273 are antigenic peptides comprising MVH Th and HIV-1 B-cell
CC epitope which may be used as a component in an anti-HIV-1 vaccine.
CC AAY91198 and AAY91199 are respectively an immunostimulatory invasin
CC protein epitope from Yersinia species, and hinge spacer peptide, both of

CC which may optionally be used in the antigenic peptides of the
CC invention.

XX Sequence 42 AA;

Query Match 69.6%; Score 94; DB 21; Length 42;
Best Local Similarity 64.0%; Pred. No. 7.7e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHPDLPREVVRSIAC 25

DB 18 cgetyqsrvtphlpralmrsttkc 42

Search completed: March 4, 2002, 12:55:50
Job time: 297 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:54:13 ; Search time 42.32 seconds
(without alignments)
44.999 Million cell updates/sec

Title: US-09-701-623C-5
Perfect score: 140
Sequence: 1 CGETYQSRVTHPLPALMRSTRKTC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 117 | 83.6 | 428 | 1 EHHU | Ig epsilon chain C |
| 2 | 113 | 80.7 | 426 | 2 I36948 | Ig epsilon-chain - |
| 3 | 80 | 57.1 | 429 | 1 EHRF | Ig epsilon chain C |
| 4 | 56 | 40.0 | 107 | 2 I68730 | IgE chain C3 regio |
| 5 | 56 | 40.0 | 107 | 2 I68726 | Ig epsilon chain C |
| 6 | 56 | 40.0 | 388 | 1 EHRM | Ig epsilon chain C |
| 7 | 56 | 40.0 | 548 | 2 S38864 | novel antigen rece |
| 8 | 54 | 38.6 | 684 | 2 S60266 | hypothetical prote |
| 9 | 52 | 37.1 | 320 | 2 C86148 | genome polyprotein |
| 10 | 50 | 35.7 | 1597 | 2 S65053 | I78K protein - tob |
| 11 | 50 | 35.7 | 1601 | 2 S48699 | insulin receptor s |
| 12 | 49.5 | 35.4 | 1235 | 1 S16948 | hypothetical prote |
| 13 | 48 | 34.3 | 38 | 2 B31194 | cellulose synthase |
| 14 | 48 | 34.3 | 1065 | 2 T52054 | protein T6D22.14 l |
| 15 | 48 | 34.3 | 2254 | 2 D86215 | gas-vesicle protel |
| 16 | 47.5 | 33.9 | 381 | 2 S28115 | gas-vesicle protel |
| 17 | 47.5 | 33.9 | 382 | 2 JQ1122 | gas-vesicle protel |
| 18 | 47.5 | 33.9 | 382 | 2 T08243 | gas-vesicle operon |
| 19 | 47 | 33.6 | 242 | 1 MF1VCJ | matrix protein M1 |
| 20 | 47 | 33.6 | 1940 | 1 S04090 | myosin heavy chain |
| 21 | 47 | 33.6 | 1940 | 1 A24922 | myosin heavy chain |
| 22 | 47 | 33.6 | 1940 | 2 A29320 | myosin heavy chain |
| 23 | 46.5 | 33.2 | 1231 | 2 S30185 | insulin receptor s |
| 24 | 46 | 32.9 | 116 | 2 S37909 | hypothetical prote |
| 25 | 46 | 32.9 | 247 | 2 A27547 | trypsin (EC 3.4.21 |
| 26 | 46 | 32.9 | 550 | 1 VGBE18 | glycoprotein E - h |
| 27 | 46 | 32.9 | 842 | 2 C83458 | conserved hypothet |
| 28 | 46 | 32.9 | 1019 | 2 A83613 | conserved hypothet |
| 29 | 45 | 32.1 | 93 | 2 I54421 | MHC RT1.B-beta2 - |

| | | | | | |
|----|------|------|------|----------|--------------------|
| 30 | 45 | 32.1 | 106 | 2 A03859 | hypothetical prote |
| 31 | 45 | 32.1 | 210 | 2 S76316 | hypothetical prote |
| 32 | 45 | 32.1 | 245 | 2 I55951 | MHC class II E-bet |
| 33 | 45 | 32.1 | 264 | 2 A60497 | H-2 class II histo |
| 34 | 45 | 32.1 | 264 | 2 S10989 | class II histocomp |
| 35 | 45 | 32.1 | 343 | 2 T32334 | hypothetical prote |
| 36 | 45 | 32.1 | 573 | 2 T21355 | hypothetical prote |
| 37 | 45 | 32.1 | 1038 | 1 MWRBCB | myosin beta heavy |
| 38 | 45 | 32.1 | 1934 | 2 I48153 | myosin heavy chain |
| 39 | 45 | 32.1 | 1935 | 1 A37102 | myosin beta heavy |
| 40 | 45 | 32.1 | 1935 | 1 S06006 | myosin beta heavy |
| 41 | 45 | 32.1 | 1935 | 2 A59286 | myosin heavy chain |
| 42 | 45 | 32.1 | 1937 | 2 I38055 | myosin heavy chain |
| 43 | 45 | 32.1 | 1938 | 1 JX0178 | talin - slime mold |
| 44 | 45 | 32.1 | 1938 | 1 JX0178 | hypothetical prote |
| 45 | 44.5 | 32.1 | 1938 | 1 JX0178 | hypothetical prote |

RESULT 1

EHHU
Ig epsilon chai
C:Species: Homo
C:Date: 31-Mar-
C:Accession: A2
R:Flanagan, J.G
EMBL J. 1, 655-
A>Title: The se
A:Reference number: A22771
A:Accession: A22771
A:Molecule type: DNA
A:Residues: 1-428 <FLA>
A:Cross-references: GB:L00022; GB:J00227; GB:V00555; NID:g185035
R:Ueda, S.; Nakai, S.; Nishida, Y.; Hisajima, H.; Honjo, T.
EMBL J. 1, 1539-1544, 1982
A>Title: Long terminal repeat-like elements flank a human immunoglobulin epsilon pseu
A:Reference number: A23195; MUID:84207910
A:Accession: A23195
A:Molecule type: DNA
A:Residues: 2-428 <UED>
A:Cross-references: GB:J00222; NID:g184755
R:Zhang, K.; Saxon, A.; Max, E.E.
J. Exp. Med. 176, 233-243, 1992

3 #text_change 16-Jul-1999
24; A94418; B93933; S02438; A53116;

epsilon heavy chain constant region 9
MUID:84236029

GB:V00555; NID:g185035

flank a human immunoglobulin epsilon pseu
MUID:84207910

encoded by alternative RNA splic
MUID:92308839

NID:g32987

R.; Sasada, R.; Igarashi, K.; Kikuchi, M.; S

sequencing of human immunoglobulin epsilon

MUID:83168897

GB:V00555; NID:g185035

R.; Leder, P.

Cell 29, 691-699, 1982

A>Title: Duplication and deletion in the human immunoglobulin epsilon genes.

MUID:83001945

A:Accession: A90824

A:Molecule type: DNA

A:Residues: 1-358, 'L', 360-428 <MAX>

A:Cross-references: GB:J00222; NID:g184755

A>Note: this sequence difference may be due to polymorphism

R:Bennich, H.H.; Johansson, S.G.O.; Von Bahr-Lindstrom, H.

in Immediate Hypersensitivity: Modern Concepts and Developments, Bach, M.K., ed., pp.

A:Reference number: A94418

A:Accession: A94418

A:Molecule type: protein

A:Residues: 'GAWTL' 6, 'X', 8-16, 'B', 18-43, 'B', 45-52, 55-92, 95-97, 'B', 99-121, 'B', 123, 'L', 124
A:Experimental source: myeloma protein Nd
R:Kenten, J.H.; Mølgaard, H.V.; Houghton, M.; Derbyshire, R.B.; Viney, J.; Bell, L.O.; G
Proc. Natl. Acad. Sci. U.S.A. 79, 6661-6665, 1982
A:Title: Cloning and sequence determination of the gene for the human immunoglobulin epsilon
A:Reference number: A93933; MUID:83065234
A:Accession: B93933
A:Molecule type: mRNA
A:Residues: 1-40; 68-114; 427-428 <KEN>
A:Cross-references: GB:L00022; NID:g185035
R:Ikeyama, S.
FEBS Lett. 224, 306-310, 1987
A:Title: Purification and characterization of a recombinant human IgE Fc-epsilon fragment
A:Reference number: S02438; MUID:88083554
A:Accession: S02438
A:Molecule type: mRNA
A:Status: nucleic acid sequence not shown
A:Residues: 98-352 <IKE>
R:Zhang, K.; Max, E.E.; Cheah, H.K.; Saxon, A.
J. Biol. Chem. 269, 456-462, 1994
A:Title: Complex alternative RNA splicing of epsilon-immunoglobulin transcripts produces
A:Reference number: A53116; MUID:94103254
A:Accession: A53116
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 320-428 <ZH2>
A:Experimental source: myeloma U266-derived cell line AF-10
A:Note: sequence extracted from NCBI backbone (NCBIN:141701, NCBIP:141702)
R:Hellman, L.
Eur. J. Immunol. 23, 159-167, 1993
A:Title: Characterization of four novel epsilon chain mRNA and a comparative analysis of
A:Reference number: A46536; MUID:93122085
A:Accession: C46536
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 392-426 <HEL>
A:Cross-references: GB:S55273; NID:g263166; PIDN:AAB24857.1; PID:g263167
A:Experimental source: B cell myeloma U-266
A:Note: sequence extracted from NCBI backbone (NCBIP:125297)
A:Accession: D46536
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 382-391 <HE2>
A:Cross-references: GB:S55276; NID:g263168; PIDN:AAB24858.1; PID:g263169
A:Experimental source: B cell myeloma U-266
A:Note: sequence extracted from NCBI backbone (NCBIP:125299)
A:Accession: A46536
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 401-428 <HE3>
A:Cross-references: GB:S53497; NID:g263162; PIDN:AAB24855.1; PID:g263163
A:Experimental source: B cell myeloma U-266
A:Note: sequence extracted from NCBI backbone (NCBIP:123483)
C:Genetics:
A:Gene: GDB:IGHE
A:Cross-references: GDB:119335; OMIM:147180
A:Map position: 14q32.33-14q32.33
A:Introns: 1/1; 104/1; 211/1; 319/1
C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (kappa) chains disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger oligomers.
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterotrimer; immunoglobulin
F:22-87/Domain: immunoglobulin homology <IM1>
F:128-195/Domain: immunoglobulin homology <IM2>
F:232-301/Domain: immunoglobulin homology <IM3>
F:338-407/Domain: immunoglobulin homology <IM4>
F:14/Disulfide bonds: interchain (to light chain) #status predicted
F:15-105, 29-89, 135-193, 239-299, 345-405/Disulfide bonds: #status predicted
F:21, 49, 99, 146, 252, 275/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:121, 209/Disulfide bonds: interchain (to heavy chain) #status predicted

Query Match

83.68; Score 117; DB 1; Length 428;

Query Match 57.1%; Score 80; DB 1; Length 429;
Best Local Similarity 60.9%; Pred. No. 0.00018;

Best Local Similarity 95.7%; Pred. No. 4.2e-10;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHLPALMRSTTK 24
||||| |||||||||
DB 294 GETYQCRVTHPHLPALMRSTTK 316

RESULT 2
I36948
Ig epsilon-chain - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C:Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C:Accession: I36948
R:Sakoyama, Y.; Hong, K.
Proc. Natl. Acad. Sci. U.S.A. 84, 1080-1084, 1987
A:Title: Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and orang
A:Reference number: I36948; MUID:87147196
A:Accession: I36948
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-426 <RES>
A:CROSS-references: GB:M15398; NID:gi176797; PIDN:AAA35416.1; PID:gi176798
C:Genetics:
A:Introns: 103/1; 209/1; 317/1
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:336-405/Domain: immunoglobulin homology <IM3>

Query Match 80.7%; Score 113; DB 2; Length 426;
Best Local Similarity 91.3%; Pred. No. 1.7e-09;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHLPALMRSTTK 24
||||| |||||||||
DB 292 GETYQCRVTHPHLPALMRSTTK 314

RESULT 3
EHRT
Ig epsilon chain C region - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 16-Jul-1999
C:Accession: A93442; A90937; A02143
R:Hellman, L.; Pettersson, U.; Engstrom, A.; Karlsson, T.; Bennlich, H.
Nucleic Acids Res. 10, 6041-6049, 1982
A:Title: Structure and evolution of the heavy chain from rat immunoglobulin E.
A:Reference number: A93442; MUID:83064537
A:Accession: A93442
A:Molecule type: mRNA
A:Residues: 1-429 <HEL>
R:Kindsvogel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.
DNA 1, 335-343, 1982
A:Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction
A:Reference number: A90937; MUID:83182019
A:Contents: myeloma IR162
A:Accession: A90937
A:Molecule type: mRNA
A:Residues: N', 169-307, 'L', 309-342 <KIN>
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:19-80/Domain: immunoglobulin homology <IM1>
F:118-186/Domain: immunoglobulin homology <IM2>
F:223-291/Domain: immunoglobulin homology <IM3>
F:327-398/Domain: immunoglobulin homology <IM4>
F:46,99,170,240,265,369,419/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 57.1%; Score 80; DB 1; Length 429;
Best Local Similarity 60.9%; Pred. No. 0.00018;

Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHLPALMRSTTK 24

I I I I I I I I I I I I I I I I

Db 284 GEGYQCRVDHPDFPKPIVRSITK 306

RESULT 4

IgE chain C3 region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000

C:Accession: I68730

R;Shinkai, Y.; Nakauchi, H.; Honjo, T.; Okumura, K.

Immunogenetics 27, 288-292, 1988

A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic acid s

A:Reference number: I54443; MUID:88152907

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F;22-90/Domain: immunoglobulin homology <IMM>

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-107 <RES>

A:Cross-references: GB:M22933; NID:g194464; PIDN:AAA37915.1; PID:g194469

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F;22-90/Domain: immunoglobulin homology <IMM>

Query Match 40.0%; Score 56; DB 2; Length 107;

Best Local Similarity 47.8%; Pred. No. 0.19;

Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHLPALMRSTTK 24

I I I I I I I I I I I I I I I I

Db 83 GYGQCIVDHPDFPKPIVRSITK 105

RESULT 5

I68726

IgE chain C3 region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000

C:Accession: I68726

R;Shinkai, Y.; Nakauchi, H.; Honjo, T.; Okumura, K.

Immunogenetics 27, 288-292, 1988

A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic acid s

A:Reference number: I54443; MUID:88152907

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F;22-90/Domain: immunoglobulin homology <IMM>

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-107 <RES>

A:Cross-references: GB:M22930; NID:g194455; PIDN:AAA37911.1; PID:g194460

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F;22-90/Domain: immunoglobulin homology <IMM>

Query Match 40.0%; Score 56; DB 2; Length 107;

Best Local Similarity 47.8%; Pred. No. 0.19;

Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHLPALMRSTTK 24

I I I I I I I I I I I I I I I I

Db 83 GYGQCIVDHPDFPKPIVRSITK 105

RESULT 6

EHMS

Ig epsilon chain C region (version 1) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 16-Jul-1999

C:Accession: A02144

R;Liu, F.T.; Albrandt, K.; Sutcliffe, J.G.; Katz, D.H.

Proc. Natl. Acad. Sci. U.S.A. 79, 7852-7856, 1982

A:Title: Cloning and nucleotide sequence of mouse immunoglobulin epsilon chain cDNA.

A:Reference number: A02144; MUID:8311774

A:Accession: A02144
A:Molecule type: mRNA
A:Residues: 1-388 <LIU>
A:Cross-references: GB:J00476; NID:g194875; PIDN:AAA38085.1; PID:g387220
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin homology
F;1-44/Domain: immunoglobulin homology (fragment) <IM1>
F;81-149/Domain: immunoglobulin homology <IM2>
F;186-254/Domain: immunoglobulin homology <IM3>
F;290-361/Domain: immunoglobulin homology <IM4>
F;10,51,62,133,205,228,332,382/Binding site: carbohydrate (Asn) (covalent) #status pr

Query Match 40.0%; Score 56; DB 1; Length 388;

Best Local Similarity 47.8%; Pred. No. 0.74;

Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHLPALMRSTTK 24

I I I I I I I I I I I I I I I I

Db 247 GYGQCIVDHPDFPKPIVRSITK 269

RESULT 7

S38864

Ig epsilon chain C region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 03-Nov-2000

C:Accession: S38864

R;Kipp, B.; Becker, W.; Schlaak, M.

submitted to the EMBL Data Library, November 1993

A:Description: Combination of a defined specificity and desired isotype by cloning of

A:Reference number: S38864

A:Accession: S38864

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-548 <KIP>

A:Cross-references: EMBL:Z27397; NID:g416537; PIDN:CAA81788.1; PID:g940782

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F;353-421/Domain: immunoglobulin homology <IMM>

Query Match 40.0%; Score 56; DB 2; Length 548;

Best Local Similarity 47.8%; Pred. No. 1.1;

Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 GETYQSRVTHPHLPALMRSTTK 24

I I I I I I I I I I I I I I I I

Db 414 GYGQCIVDHPDFPKPIVRSITK 436

RESULT 8

S60266

novel antigen receptor precursor - nurse shark

C:Species: Ginglymostoma cirratum (nurse shark)

C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 21-Jul-2000

C:Accession: S60266

R;Greenberg, A.S.; Avila, D.; Hughes, M.; Hughes, A.; McKinney, E.C.; Flajnik, M.F.

Nature 374, 168-173, 1995

A:Title: A new antigen receptor gene family that undergoes rearrangement and extensiv

A:Reference number: S60266; MUID:95183140

A:Accession: S60266

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-684 <GRE>

A:Cross-references: EMBL:U18701; NID:g699442; PIDN:AAB48195.1; PID:g699443

Query Match 38.6%; Score 54; DB 2; Length 684;

Best Local Similarity 57.8%; Pred. No. 2.7;

Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

Qy 4 TYQSRVTHPHLPALMRSTTK 24
 Db 642 TYSCLVGHPSLRDLIRSTNK 662

RESULT 9
C86148
hypothetical protein AAF78401.1 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001; #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C86148
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huizlar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: C86148
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-320 <STO>
A:Cross-references: GB:AE005172; NID:g8671838; PIDN:AAF78401.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Superfamily: acyl-CoA thioesterase II

| | | | | |
|-----------------------|-------|----------------|-------|-------------|
| Query Match | 37.1% | Score 52: | DB 2: | Length 320: |
| Best Local Similarity | 60.0% | Pred. No. 2.5: | | |
| Matches | 9: | Conservative | 3: | Mismatches |
| | | | 3: | Indels |
| | | | 0: | Gaps |
| | | | 0: | Gaps |

QY 3 ETYQSRVTHPLPRA 17
| :||:| ||||| :
DB 142 ELRESRITDPLPRS 156

RESULT 10
S65053 genome polyprotein - Chinese rape mosaic virus
N:Alternate names: RNA replicase
N:Contents: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: Chinese rape mosaic virus
C>Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 17-Mar-1999
C:Accession: S65053
A:Author: R.Aguilar, I.; Sanchez, F.; Martin Martin, A.; Martinez-Herrera, D.; Ponz, F.
A:Title: Nucleotide sequence of Chinese rape mosaic virus (oilseed rape mosaic virus), a
A:Reference number: S65053; MUID:96197410
A:Accession: S65053
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: genomic RNA
A:Residues: 1-1597 <AGU>
A:Cross-references: EMBL:U30944
A>Note: readthrough of the terminator UGA occurs between codons CAA for 1103-Gln and CAA
A>Note: the internal stop codon is translated as X
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1995
C:Superfamily: cucumber mosaic virus RNA 1 protein
C:Keywords: nucleotidyltransferase

| | | | | |
|-----------------------|-----------------|---------------|------------|--------------|
| Query Match | 35.7% | Score 50; | DB 2; | Length 1597; |
| Best Local Similarity | 36.8% | Pred. No. 26; | | |
| Matches 14; | Conservative 4; | Mismatches 6; | Indels 14; | Gaps 2; |

QY 2 GETYQS---RVT-----HPHLPALMRSTTKC 25
||||| :||| ||: || | ||:|
Db 1031 GETYEKTAIVLTTATPLEIISRASPHVLVALTRHTTTC 1068

```

RESULT 11
S48699
178K protein - tobacco mosaic virus (strain cr-TMW)
N:Contains: 122K protein
N:Alternate names: readthrough protein
C:Species: tobacco mosaic virus, TMW
A:Variety: strain cr-TMW
C:Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 20-Sep-1999
C:Accession: S48699; S48659
R:Dorokhov, Y.L.; Ivanov, P.A.; Novikov, V.K.; Agranovsky, A.A.; Morozov, S.Y.; Efimo
FEBS Lett. 350, 5-8, 1994
A:Title: Complete nucleotide sequence and genome organization of a tobamovirus infect
A:Reference number: S48659; MUID:94341372
A:Accession: S48699
A:Molecule type: genomic RNA
A:Residues: 1-1601 <DOR>
A:Cross-references: EMBL:229370; NID:g488713; PIDN:CAA82559.1; PID:g619908
A:Experimental source: tobamovirus infecting cruciferae plants (cr-TMW)
A:Note: readthrough of the terminator UGA occurs between codons CAA for 1107-Gln and
A:Note: the internal stop codon is translated as X
C:Superfamily: cucumber mosaic virus RNA 1 protein
F:1-1601/Product: 178K protein #status predicted <PRO2>
F:1-1107/Product: 122K protein #status predicted <PRO1>

```

| | | | | |
|-----------------------|--------------|---------------|------------|--------------|
| Query Match | 35.7% | Score 50; | DB 2; | Length 1601; |
| Best Local Similarity | 36.8% | Pred. No. 26; | | |
| Matches 14; | Conservative | 4; | Mismatches | 6; |
| | Indels | 14; | Gaps | 2; |

Qy 2 GETYQS---RVT-----HPHLPRALMRSTTKC 25
||||| :|: ||| |

Dd 1035 GETVEKTAIVRLTSTPLEIISRASHVLVALTRHTTRC 1072

RESULT 12

insulin receptor substrate IRS-1 - rat
S16948

N:Alternate names: insulin receptor substrate pp185
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C:Accession: S16948; A39811
R:Sun, X.J.; Rothenberg, P.; Kahn, C.R.; Backer, J.M.; Araki, E.; Wiiden, P.A.; Cahill
Nature 352, 73-77, 1991

A:title: Structure of the insulin receptor substrate IRS-1 defines a unique signal tr
A:Reference number: S16948; MUID:91287824

A:Accession: S16948

A:Molecule type: mRNA

A:Residues: 1-1235 <SUN>

A:Cross-references: EMBL:X58375; NID:g56503; PIDN:CAA41264.1; PID:g56504
R:Rothenberg, P.L.; Lane, W.S.; Karasik, A.; Backer, J.; White, M.; Kahn, C.R.
J. Biol. Chem. 266, 8302-8311, 1991

A:title: Purification and partial sequence analysis of ppl85, the major cellular subs
A:Reference number: A39811; MUID:91217066

A:Accession: A39811

A:Molecule type: Protein

A:Residues: 44-51;173-178;223-243;489-506;635-646;871-882,'I',884,'X',886-888;932-936

A:Note: the phosphotyrosine residue was not identified

C:Comment: This protein and the beta chain of the insulin receptor itself are the maj
C:Comment: Phosphorylation of this protein in response to insulin is maximal at 30 se
insulin.

C:Superfamily: insulin receptor substrate IRS-1; pleckstrin repeat homology

C:Keywords: phosphoprotein; signal transduction

F:11-113/domain: pleckstrin repeat homology <PLK>

F:872-891/Region: glutamine-rich

| | | | | |
|-----------------------|--------------|-------------|------------|--------------|
| Query Match | 35.4%; | Score 49.5; | DB 1; | Length 1235; |
| Best Local Similarity | 47.8%; | | | |
| Pred. No. 24; | | | | |
| Matches 11; | Conservative | 1; | Mismatches | 4; |
| Indels | 7; | Gaps | 1; | |

```

QY      1 CGETYQSRVTH-----PHLPR 16
      ||      : |||      |||||

```

Db 816 CGARPESVTHPHHHALOPHLPR 838

RESULT 13

B31194
B31194 10
Hypothetical protein ORF1 (FGF 5' region) - human
C:Species: Homo sapiens (man)
C>Date: 12-Jan-1989 #sequence_revision 12-Jan-1989 #text_change 05-Nov-1999
C;Accession: B31194
R;Zhan, X.; Bates, B.; Hu, X.; Goldfarb, M.
Mol. Cell. Biol. 8, 3487-3495, 1988
A>Title: The human FGF-5 oncogene encodes a novel protein related to fibroblast growth factor
A;Reference number: A31194; MUID:89096942
A;Accession: B31194
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A;Residues: 1-38 <ZHA>
A;Cross-references: GB:M21617; NID:c8006636; PID:AAB60698.1; PID:c182541

Query Match 34.3%; Score 48; DB 2; Length 38;
Best Local Similarity 69.2%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGETYQSRVTHPH 13
Db 5 CGEAGOSRGTOPH 17

RESULT 14

T52054
cellulose synthase (EC 2.4.1.-) catalytic subunit [validated] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 31-Dec-2000
C:Accession: T52054
R:Arioli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Camilleri, C.; Herth, W.;
Science 279, 717-720, 1998
A:title: Molecular analysis of cellulose biosynthesis in Arabidopsis.
A:reference number: Z13745; MUID:98111412
A:Accession: T52054
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1065 <ARI>
A:cross-references: EMBL:AF021714; PIDN:AAC393336.1
C:Genetics:
A:Note: Ath-B
C:Function:
A:description: EC 2.4.1.-; cellulose synthase [validated; MUID:98111412]; involved in as
C:keywords: glycosyltransferase; hexosyltransferase

Query Match 34.3%; Score 48; DB 2; Length 1065;
Best Local Similarity 47.4%; Pred. No. 35;
Matches 9; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

| | | | |
|----|-----|---------------------|-----|
| QY | 5 | YQSRVTHPHLPALMRSTT | 23 |
| | | : | |
| Db | 125 | YDKEVSHNHLPRLTSRQDT | 143 |

RESULT 15

protein T6b22.14 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D86215
R:Theologians, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:58:03 ; Search time 24.65 Seconds
(without alignments)
37.185 Million cell updates/sec

Title: US-09-701-623C-5
Perfect score: 140
Sequence: 1 CGETYQSRVTHPLRALMRSTTKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|-------|-------|--------|--------------|--------------------|
| 1 | 117 | 83.6 | 428 | 1 EPC_HUMAN | P01854 homo sapien |
| 2 | 80 | 57.1 | 429 | 1 EPC_RAT | P01855 rattus norv |
| 3 | 56 | 40.0 | 421 | 1 EPC_MOUSE | P06336 mus musculu |
| 4 | 50 | 35.7 | 1597 | 1 RPPO_CRMV | Q66220 chinese rap |
| 5 | 49.5 | 35.4 | 1235 | 1 IRS1_RAT | P35570 rattus norv |
| 6 | 47.5 | 33.9 | 381 | 1 GVPC_HALME | Q02228 halobacteri |
| 7 | 47.5 | 33.9 | 382 | 1 GVCL_HALN1 | P24574 halobacteri |
| 8 | 47 | 33.6 | 242 | 1 VMAT_INCJJ | P12446 influenza c |
| 9 | 47 | 33.6 | 1938 | 1 MYHD_HUMAN | Q9UKX3 homo sapien |
| 10 | 47 | 33.6 | 1940 | 1 MYH3_CHICK | P02565 gallus gall |
| 11 | 47 | 33.6 | 1940 | 1 MYH3_HUMAN | P11055 homo sapien |
| 12 | 47 | 33.6 | 1940 | 1 MYH3_RAT | P12847 rattus norv |
| 13 | 46.5 | 33.2 | 1233 | 1 IRS1_MOUSE | P35569 mus musculu |
| 14 | 46 | 32.9 | 116 | 1 YK14_YEAST | P36078 saccharomyc |
| 15 | 46 | 32.9 | 247 | 1 TRY3_RAT | P08426 rattus norv |
| 16 | 46 | 32.9 | 550 | 1 VGLL_HSV11 | P04488 herpes simp |
| 17 | 45 | 32.1 | 106 | 1 Y116_ADE02 | P03287 human adeno |
| 18 | 45 | 32.1 | 264 | 1 HB2D_RAT | P18211 rattus norv |
| 19 | 45 | 32.1 | 736 | 1 MYH7_RABIT | P04461 oryctolagus |
| 20 | 45 | 32.1 | 777 | 1 BAR1_HUMAN | Q99728 homo sapien |
| 21 | 45 | 32.1 | 1601 | 1 RPPO_TVCV | Q88920 turnip vein |
| 22 | 45 | 32.1 | 1934 | 1 MYH7_MESAU | P13540 mesocricetu |
| 23 | 45 | 32.1 | 1935 | 1 MYH7_HUMAN | P12883 homo sapien |
| 24 | 45 | 32.1 | 1935 | 1 MYH7_PIG | P79293 sus scrofa |
| 25 | 45 | 32.1 | 1935 | 1 MYH7_RAT | P02564 rattus norv |
| 26 | 45 | 32.1 | 1937 | 1 MYH8_HUMAN | P13535 homo sapien |
| 27 | 45 | 32.1 | 1938 | 1 MYSS_CHICK | P13538 gallus gall |
| 28 | 45 | 32.1 | 1939 | 1 MYH1_HUMAN | P12882 homo sapien |
| 29 | 45 | 32.1 | 1939 | 1 MYH4_HUMAN | Q9Y623 homo sapien |
| 30 | 45 | 32.1 | 1941 | 1 MYH2_HUMAN | Q9UKX2 homo sapien |
| 31 | 45 | 32.1 | 2492 | 1 TALA_DICDI | P54633 dictyosteli |
| 32 | 44.5 | 31.8 | 450 | 1 MUC_CANFA | P01874 canis fami |
| 33 | 44 | 31.4 | 340 | 1 ALC2_HUMAN | P01877 homo sapien |

ALIGNMENTS

RESULT_1

| ID | EPC_HUMAN | STANDARD; | PRT; | 428 AA. |
|----|------------------------------------------------------------------------|-----------|------|---------|
| AC | P01854; | | | |
| DT | 21-JUL-1986 (Rel. 01, Created) | | | |
| DT | 21-JUL-1986 (Rel. 01, Last sequence update) | | | |
| DT | 20-AUG-2001 (Rel. 40, Last annotation update) | | | |
| DE | IG EPSILON CHAIN C REGION. | | | |
| GN | IGHE. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=83168897; PubMed=6300763; | | | |
| RA | Seno M., Kurokawa T., Ono Y., Onda H., Sasada R., Igarashi K., | | | |
| RA | Kikuchi M., Sugino Y., Nishida Y., Honjo T.; | | | |
| RT | "Molecular cloning and nucleotide sequencing of human immunoglobulin | | | |
| RT | epsilon chain cDNA."; | | | |
| RL | Nucleic Acids Res. 11:719-726(1983). | | | |
| RN | [2] | | | |
| RX | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=63001945; PubMed=6288268; | | | |
| RA | Max E.E., Battey J., Ney R., Kirsch I.R., Leder P.; | | | |
| RT | "Duplication and deletion in the human immunoglobulin epsilon genes."; | | | |
| RN | [3] | | | |
| RX | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=84236029; PubMed=6234164; | | | |
| RA | Flanagan J.G., Rabbitts T.H.; | | | |
| RT | "The sequence of a human immunoglobulin epsilon heavy chain constant | | | |
| RT | region gene, and evidence for three non-allelic genes."; | | | |
| RL | EMBO J. 1:655-660(1982). | | | |
| RN | [4] | | | |
| RX | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=84207910; PubMed=6327276; | | | |
| RA | Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.; | | | |
| RT | "Long terminal repeat-like elements flank a human immunoglobulin | | | |
| RT | epsilon pseudogene that lacks introns."; | | | |
| RL | EMBO J. 1:1539-1544(1982). | | | |
| RN | [5] | | | |
| RX | PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND). | | | |
| RA | Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.; | | | |
| RL | (In) Bach M.K. (eds.); | | | |
| RL | Immediate hypersensitivity: modern concepts and developments, pp.1-36, | | | |
| RL | Marcel Dekker, New York (1978). | | | |
| RN | [6] | | | |
| RX | SEQUENCE OF 1-40: 68-114 AND 427-428 FROM N.A. | | | |
| RX | MEDLINE=83065234; PubMed=6815656; | | | |
| RA | Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J., | | | |
| RA | Bell L.O., Gould H.J.; | | | |
| RT | "Cloning and sequence determination of the gene for the human | | | |
| RT | immunoglobulin epsilon chain expressed in a myeloma cell line."; | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982). | | | |
| RN | [7] | | | |

| | | | | | | |
|----|------|------|------|---|------------|--------------------|
| 34 | 44 | 31.4 | 345 | 1 | HALF_CHICK | P15979 gallus gall |
| 35 | 44 | 31.4 | 425 | 1 | HE47_CAEEL | Q18212 caenorhabdi |
| 36 | 44 | 31.4 | 427 | 1 | HE47_PIG | Q29024 sus scrofa |
| 37 | 44 | 31.4 | 428 | 1 | HE47_HUMAN | Q13838 homo sapien |
| 38 | 44 | 31.4 | 428 | 1 | HE47_RAT | Q63413 rattus norv |
| 39 | 44 | 31.4 | 465 | 1 | OPCA_NOSPU | P48971 nostoc punc |
| 40 | 44 | 31.4 | 593 | 1 | COX1_HALHA | P33518 halobacteri |
| 41 | 44 | 31.4 | 1075 | 1 | NFC3_HUMAN | Q12968 homo sapien |
| 42 | 43.5 | 31.1 | 454 | 1 | MUC_MESAU | P06337 mesocricetu |
| 43 | 43.5 | 31.1 | 557 | 1 | C791_SORBI | Q43135 sorghum bic |
| 44 | 43 | 30.7 | 80 | 1 | RL31_MYCTU | Q10608 mycobacteri |
| 45 | 43 | 30.7 | 84 | 1 | RL31_MYCLE | P45834 mycobacteri |


```
RL Submitted (APR-1986) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 34-421 FROM N.A.
RX MEDLINE-83117774; PubMed-6818553;
RA Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;
RT "Cloning and nucleotide sequence of mouse immunoglobulin epsilon
chain cDNA."
RL Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856(1982).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01857; CAA25977.1; -
DR EMBL; X01857; CAA25978.1; -
DR PIR; A02145; EHMSS.
DR PIR; A02144; EHMS.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; Igcl; 2.
DR SMART; SM00410; Ig_like; 2.
DR PROSITE; PS00290; Ig_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 90 CH1.
FT DOMAIN 91 197 CH2.
FT DOMAIN 198 304 CH3.
FT DOMAIN 305 421 CH4.
FT DISULFID 23 75 BY SIMILARITY.
FT DISULFID 121 180 BY SIMILARITY.
FT DISULFID 226 285 BY SIMILARITY.
FT DISULFID 330 392 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE 421 AA; 47320 MW; 8F909E1F30A06B47 CRC64;

Query Match 40.0%; Score 56; DB 1; Length 421;
Best Local Similarity 47.8%; Pred. No. 0.29;
Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 GETVQSRVTHPLPRALMRSTTK 24
| | | | | | | | | | | | | |
Db 280 GYGQCIVDHPDPKPIVRSTK 302

RESULT 4
RRPO_CRMV STANDARD; PRT; 1597 AA.
ID RPO_CRMV
AC Q66220;
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.48) (182 KDA PROTEIN) [CONTAINS:
DE METHYLTRANSFERASE/RNA HELICASE (MT/HEL) (125 KDA PROTEIN)].
OS Chinese rape mosaic virus (CRMV) (Oilseed rape mosaic virus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OX NCBI_TaxID=42007;
RN SEQUENCE FROM N.A.

RX MEDLINE-96197410; PubMed-8616237;
RA Aguilar I., Sanchez F., Martin-Martin A., Martinez-Herrera D.,
RA Ponz F.;
RT "Nucleotide sequence of Chinese rape mosaic virus (oilseed rape mosaic
virus), a crucifer tobamovirus infections on Arabidopsis thaliana.";
PLant Mol. Biol. 30:191-197(1996).
CC -/- FUNCTION: THE LONGER PROTEIN IS A RNA POLYMERASE ACTIVE IN VIRAL
CC RNA REPLICATION.
CC -/- FUNCTION: THE SMALLER PROTEIN IS A METHYLTRANSFERASE ACTIVE IN RNA
CC CAPPING AND AN RNA HELICASE.
CC -/- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR CODON UAG OCCURS
CC BETWEEN CODONS FOR GLN-1103 AND GLN-1105.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U30944; AAB60599.1; -
DR InterPro; IPR001788; RNA_dep_RNAPol2.
DR InterPro; IPR002588; V_methyltransf.
DR InterPro; IPR000606; Viral_helicaseI.
DR Pfam; PF00978; RNA_dep_RNAPol2; 1.
DR Pfam; PF01443; Viral_helicaseI; 1.
DR Pfam; PF01660; Vmethyltransf; 1.
KW Transferase; RNA-directed RNA polymerase; Helicase; ATP-binding.
FT CHAIN 1 1597 RNA-DIRECTED RNA POLYMERASE.
FT CHAIN 1 1103 METHYLTRANSFERASE/RNA HELICASE.
FT NP_BIND 823 830 ATP (POTENTIAL).
FT VARIANT 919 919 V -> G.
FT VARIANT 1286 1286 H -> Y.
FT SEQUENCE 1597 AA; 181621 MW; 7256A908BD3308F2 CRC64;

Query Match 35.7%; Score 50; DB 1; Length 1597;
Best Local Similarity 36.8%; Pred. No. 10;
Matches 14; Conservative 4; Mismatches 6; Indels 14; Gaps 2;

QY 2 GETVQS---RVT-----HPLPLRALMRSTTK 25
| | | | | | | | | | | | | |
Db 1031 GETYKTAIVRLTATPLEIISRASPHVLTRHTTC 1068

RESULT 5
IRSL_RAT STANDARD; PRT; 1235 AA.
ID IRS1_RAT
AC P35570;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE INSULIN RECEPTOR SUBSTRATE-1.
GN IRS1 OR IRS-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Eukarya; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RX MEDLINE-91287824; PubMed=1648180;
RA Sun X.J., Rothenberg P., Kahn C.R., Backer J.M., Araki E.,
RA Wilden P.A., Cahill D.A., Goldstein B.J., White M.F.;
RT "Structure of the insulin receptor substrate IRS-1 defines a unique
RT signal transduction protein.";
RL Nature 352:73-77(1991).
RN [2]
RP PHOSPHORYLATION SITES.
RX MEDLINE-94067102; PubMed=7504175;
RA Sun X.J., Crimmins D.L., Myers M.G., Miralpeix M., White M.F.;
RT "Pleiotropic insulin signals are engaged by multisite phosphorylation
```

```

of IRS-1."
RL Mol. Cell. Biol. 13:7418-7428(1993).
RN [3]
RP PHOSPHORYLATION SITES.
RX MEDLINE-93352637; PubMed-9349691;
RA Tanasijevic M.J., Myers M.G., Thoma R.S., Crimmins D.L., White M.F.,
RT Sacks D.B.;
RT "Phosphorylation of the insulin receptor substrate IRS-1 by casein
kinase II."
RL J. Biol. Chem. 268:18157-18166(1993).
CC -1- FUNCTION: MAY MEDATE THE CONTROL OF VARIOUS CELLULAR PROCESSES
CC BY INSULIN. WHEN PHOSPHORYLATED BY THE INSULIN RECEPTOR BINDS
CC SPECIFICALLY TO VARIOUS CELLULAR PROTEINS CONTAINING SH2 DOMAINS
CC SUCH AS PHOSPHATIDYLINOSITOL 3-KINASE P85 SUBUNIT OR GRB-2.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PTB DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X58375; CAA41264.1; -.
DR PIR; S16948; S16948.
DR HSP; P35568; IIRS.
DR InterPro; IPR002404; Insulin_Recep_S-1.
DR InterPro; IPR001849; PH.
DR Pfam; PF02174; IRS; 1.
DR PRINTS; PF00169; PH; 1.
DR PRINTS; P00628; INSULINRSI.
DR SMART; SM00233; PH; 1.
DR SMART; SM00310; PTBI; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
KW PHOSPHORYLATION.
FT DOMAIN 12 115 PH.
FT MOD_RES 152 262 PTB.
FT DOMAIN 152 262 POLY-GLN.
FT DOMAIN 172 881 POLY-GLN.
FT DOMAIN 1196 1200 POLY-PRO.
FT MOD_RES 99 99 PHOSPHORYLATION (BY CK2).
FT MOD_RES 460 460 PHOSPHORYLATION (BY INSR).
FT MOD_RES 502 502 PHOSPHORYLATION (BY CK2).
FT MOD_RES 508 508 PHOSPHORYLATION (BY INSR).
FT MOD_RES 528 628 PHOSPHORYLATION (BY INSR).
FT MOD_RES 895 895 PHOSPHORYLATION (BY INSR).
FT MOD_RES 939 939 PHOSPHORYLATION (BY INSR).
FT MOD_RES 987 987 PHOSPHORYLATION (BY INSR).
FT MOD_RES 1172 1172 PHOSPHORYLATION (BY INSR).
FT MOD_RES 1222 1222 PHOSPHORYLATION (BY INSR).
SQ SEQUENCE 1235 AA; 131178 MW; A274BC7540CA85C5 CRC64;

Query Match 35.4%; Score 49.5; DB 1; Length 1235;
Best Local Similarity 47.8%; Pred. No. 9.1;
Matches 11; Conservative 1; Mismatches 4; Indels 7; Gaps 1;

QY 1 CGETYQSRVTH-----PHLPR 16
||| :| |||
Db 816 CGARPESSVTHPHHHPHALQPHLPR 838

RESULT 6
GVPC_HALME STANDARD; PRT; 381 AA.
ID GVPC_HALME
AC Q02228;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE GAS VESICLE PROTEIN C.
GN GVPC.
OS Halobacterium mediterranei (Haloflex mediterranei).

```

```

Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloflex.
NCBI_TaxID=2252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1411;
RX MEDLINE-93021102; PubMed-1404376;
RA Englert C., Krueger K., Offner S., Pfeifer F.;
RT "Three different but related gene clusters encoding gas vesicles in
halophilic archaea."
RL J. Mol. Biol. 227:586-592(1992).
CC -1- FUNCTION: MAY CONFERS STABILITY TO THE GAS VESICLE MEMBRANES. GAS
CC VESICLES ARE SMALL, HOLLOW, GAS FILLED PROTEIN STRUCTURES THAT ARE
CC FOUND IN SEVERAL PROKARYOTIC PLANKTONIC MICROORGANISMS. THEY ALLOW
CC THE POSITIONING OF THE BACTERIA, AT THE FAVORABLE DEPTH FOR
CC GROWTH.
CC -1- SUBCELLULAR LOCATION: BINDS TO THE EXTERNAL SURFACE OF THE GAS
CC VESICLE MEMBRANE.
CC -1- SIMILARITY: VERY DISTANT, TO GAS VESICLE PROTEINS TYPE C OF
CC CYANOBACTERIA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X64701; CAA45944.1; -.
DR PIR; S28115; S28115.
DR Gas vesicle; Repeat.
KW DOMAIN 22 274 7 X APPROXIMATE TANDEM REPEATS.
FT REPEAT 22 59 1.
FT REPEAT 60 84 2.
FT REPEAT 85 122 3.
FT REPEAT 123 160 4.
FT REPEAT 161 192 5.
FT REPEAT 193 232 6.
FT REPEAT 233 274 7.
FT DOMAIN 246 335 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 381 AA; 42653 MW; 9FB48199D0305921 CRC64;

Query Match 33.9%; Score 47.5; DB 1; Length 381;
Best Local Similarity 64.3%; Pred. No. 5.3;
Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 CGETYQSRVTHPHL 14
||| ||: |||
Db 343 CGEYQQA-ITEPHL 355

RESULT 7
GVCL_HALNI
ID GVCL_HALNI STANDARD; PRT; 382 AA.
AC P24574; Q9H117;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GAS VESICLE PROTEIN C 1.
GN (GVPC1 OR GVPC OR VNG5032G) AND (GVPC2 OR VNG6031G).
OS Halobacterium sp. (strain NRC-1), and
OS Halobacterium sp. (strain NRC-817).
OG Plasmid pNRC100, plasmid pNRC200, and plasmid pH1.
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC Halobacterium.
OX NCBI_TaxID=64091, 148370;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRC-1; PLASMID=pNRC100;
RX MEDLINE-91323716; PubMed-1864501;
RA Jones J.G., Young D.C., Dassarma S.;
RT "Structure and organization of the gas vesicle gene cluster on the

```

RT Halobacterium halobium plasmid pNRC100.";
 RL Gene 102:117-122(1991).
 RN [2]
 RC SEQUENCE FROM N.A.
 RX STRAIN=NRC-1; PLASMID=pNRC100;
 RA MEDLINE=99063795; PubMed=9847077;
 NG W.V., Clufo S.A., Smith T.M., Bungarner R.E., Baskin D., Faust J.,
 RA Hall B., Loretz C., Seto J., Stigel J., Hood L., Dassarma S.;
 RT "Snapshot of a large dynamic replicon in a halophilic archaeon:
 RT megaplasmid or minichromosome?";
 RL Genome Res. 8:1131-1141(1998).
 RN [3]
 RN SEQUENCE FROM N.A.
 RP STRAIN=NRC-1; PLASMID=pNRC200;
 RC MEDLINE=20504483; PubMed=11016950;
 RX Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 RT "Genome sequence of Halobacterium species NRC-1";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 RN [4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=NRC-817; PLASMID=PHH1;
 RC MEDLINE=92065812; PubMed=1956294;
 RX Horne M., Englert C., Wimmer C., Pfeifer F.;
 RA "A DNA region of 9 kbp contains all genes necessary for gas vesicle
 RT synthesis in halophilic archaeobacteria";
 RL Mol. Microbiol. 5:1159-1174(1991).
 RN [5]
 RN SEQUENCE FROM N.A.
 RP STRAIN=NRC-817; PLASMID=PHH1;
 RC MEDLINE=93021102; PubMed=1404376;
 RX Englert C., Krueger K., Offner S., Pfeifer F.;
 RA "Three different but related gene clusters encoding gas vesicles in
 RT halophilic archaea";
 RL J. Mol. Biol. 227:586-592(1992).
 CC -!- FUNCTION: MAY CONFER STABILITY TO THE GAS VESICLE MEMBRANES. GAS
 CC VESICLES ARE SMALL, HOLLOW, GAS FILLED PROTEIN STRUCTURES THAT ARE
 CC FOUND IN SEVERAL PROKARYOTIC PLANKTONIC MICROORGANISMS. THEY ALLOW
 CC THE POSITIONING OF THE BACTERIA, AT THE FAVORABLE DEPTH FOR
 CC GROWTH.
 CC -!- SUBCELLULAR LOCATION: BINDS TO THE EXTERNAL SURFACE OF THE GAS
 CC VESICLE MEMBRANE.
 CC -!- SIMILARITY: VERY DISTANT, TO GAS VESICLE PROTEINS TYPE C OF
 CC CYANOBACTERIA.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M58557; AAA98197.1; -;
 CC EMBL; AF016485; AAC82810.1; -;
 CC EMBL; AE005142; AAG20727.1; -;
 CC EMBL; X57161; CAA40451.1; -;
 CC EMBL; X64729; CAA45981.1; -;
 CC PIR; J01122; J01122;
 CC PIR; S15483; S15483;
 CC PIR; S28128; S28128;
 CC Gas vesicle; Plasmid; Repeat; Complete proteome.
 KW DOMAIN 22 284 7 X APPROXIMATE TANDEM REPEATS.
 FT REPEAT 22 60 1.
 FT REPEAT 61 92 2.
 FT REPEAT 93 130 3.
 FT REPEAT 131 168 4.

FT REPEAT 169 200 5.
 FT REPEAT 201 240 6.
 FT REPEAT 241 284 7.
 FT DOMAIN 254 336 ASP/GLU-RICH (ACIDIC).
 SQ SEQUENCE 382 AA; 42391 MW; 171DBEB4C0364F46 CRC64;
 Query Match 33.9%; Score 47.5; DB 1; Length 382;
 Best Local Similarity 64.3%; Pred. No. 5.3;
 Matches 9; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
 QY 1 CGETYQSRVTHPHL 14
 Db 344 CGEYQA-ITEPHL 356
 RESULT 8
 VMAT_INCUJ STANDARD; PRT; 242 AA.
 ID VMAT_INCUJ
 AC P12446;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE MATRIX PROTEIN.
 GN M.
 OS Influenza C virus (strain C/JJ/50).
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 OC Influenza C virus.
 OC NCBI_TaxID=11560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88300888; PubMed=3404579;
 RA Yamashita M., Krystal M., Palese P.;
 RT "Evidence that the matrix protein of influenza C virus is coded for
 RL by a spliced mRNA";
 RL J. Virol. 62:3348-3355(1988).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M22038; AAA43781.1; -;
 CC PIR; A28878; MFIVCJ.
 KW Matrix protein.
 SQ SEQUENCE 242 AA; 26967 MW; ABF9D9054E1C9D91 CRC64;
 Query Match 33.6%; Score 47; DB 1; Length 242;
 Best Local Similarity 38.9%; Pred. No. 3.9;
 Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 QY 1 CGETYQSRVTHPHLPRAL 18
 Db 209 CNHTFGSNIMRPHLEKAI 226
 RESULT 9
 MYHD_HUMAN STANDARD; PRT; 1938 AA.
 ID MYHD_HUMAN
 AC Q9UKX3; O95252;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE MYOSIN HEAVY CHAIN, SKELETAL MUSCLE, EXTRAOCULAR (MYHC-BO).
 GN MYH13.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;


```

DR PRINTS; PD00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family.
FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
FT DOMAIN 786 815 IQ.
FT DOMAIN 844 1940 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP.
FT DOMAIN 660 682 ACTIN-BINDING.
FT DOMAIN 762 776 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 700 700 ALKYLATION (SH-1).
FT MOD_RES 710 710 ALKYLATION (SH-2).
FT VARIANT 379 379 G -> D.
FT CONFLICT 1547 1547 T -> A (IN REF. 2).
FT CONFLICT 1913 1915 ERA -> GRT (IN REF. 2).
SQ SEQUENCE 1940 AA; 222816 MW; C348333D75B04DF2 CRC64;

Query Match 33.6%; Score 47; DB 1; Length 1940;
Best Local Similarity 52.9%; Pred. No. 36;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 8 RVTHPLRALMRSTK 24
| | | | | | | | | |
Db 668 RSTHPHFVRLIPNETK 684

RESULT 11
MYH3.HUMAN
ID MYH3.HUMAN STANDARD; PRT; 1940 AA.
AC P11055; Q15492;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOSIN HEAVY CHAIN, FAST SKELETAL MUSCLE, EMBRYONIC (MUSCLE EMBRYONIC
DE MYOSIN HEAVY CHAIN) (SMCE).
GN MYH3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263803; PubMed=2726495;
RA Eller M.S., Stedman H.H., Sylvester J.E., Fertels S.H.,
RA Rubinstein N.A., Kelly A.M., Sarkar S.;
RT "Nucleotide sequence of full length human embryonic myosin heavy
RT chain cDNA.";
RL Nucleic Acids Res. 17:3591-3592(1989).
RN [2]
RP SEQUENCE OF 774-1940 FROM N.A.
RX MEDLINE=90033298; PubMed=2806546;
RA Eller M.S., Stedman H.H., Sylvester J.E., Fertels S.H., Wu Q.-L.,
RA Raychowdhury M.K., Rubinstein N.A., Kelly A.M., Sarkar S.;
RT "Human embryonic myosin heavy chain cDNA. Interspecies sequence
RT conservation of the myosin rod, chromosomal locus and isoform
RT specific transcription of the gene.";
RL FEBS Lett. 256:21-28(1989).
RN [3]
RP SEQUENCE OF 856-1940 FROM N.A.
RC TISSUE-Skeletal muscle;
RX MEDLINE=90235862; PubMed=1691980;
RA Bober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,
RA Arnold H.H.;
RT "Identification of three developmentally controlled isoforms of human
RT myosin heavy chains.";
RL Eur. J. Biochem. 189:55-65(1990).
RN [4]

SEQUENCE OF 856-1940 FROM N.A.
MEDLINE=89366648; PubMed=2771643;
Karsch-Mizrachi I., Travis M., Blau H., Leinwand L.A.;
"Expression and DNA sequence analysis of a human embryonic skeletal
muscle myosin heavy chain gene.";
Nucleic Acids Res. 17:6167-6179(1989).
FUNCTION: MUSCLE CONTRACTION.
SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
AND 2 REGULATORY LIGHT CHAIN SUBUNITS (RLC-2).
SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
DEVELOPMENTAL STAGE: ABUNDANTLY PRESENT IN FETAL SKELETAL MUSCLE
AND NOT PRESENT OR BARELY DETECTABLE IN HEART AND ADULT SKELETAL
MUSCLE.
DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
SUBFRAGMENT (S2).
SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
SIMILARITY: CONTAINS 1 IQ DOMAIN.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; X13988; CAA32167.1;
EMBL; X13100; CAA31492.1;
EMBL; X51593; CAA35942.1;
EMBL; X15696; CAA33731.1;
PIR; S04090; S04090.
HSP; P13538; 2MYS.
MIM; 160720;
InterPro: IPR000048; IQ.
InterPro: IPR002928; Myosin_tail.
InterPro: IPR001609; myosin_head.
Pfam: PF00612; IQ; 1.
Pfam: PF00063; myosin_head; 1.
Pfam: PF01576; Myosin_tail; 1.
PRINTS; PD00193; MYOSINHEAVY.
ProDom; PD000355; myosin_head; 1.
SMART; SM00015; IQ; 1.
SMART; SM00242; MYSC; 1.
PROSITE; PS50096; IQ; 1.
Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
Calmodulin-binding; ATP-binding; Methylation; Alkylation;
Multigene family.
DOMAIN 1 781 MYOSIN HEAD-LIKE.
DOMAIN 782 811 IQ.
DOMAIN 840 1933 COILED COIL (POTENTIAL).
NP_BIND 179 186 ATP (POTENTIAL).
DOMAIN 656 678 ACTIN-BINDING.
DOMAIN 758 772 ACTIN-BINDING.
MOD_RES 130 130 METHYLATION (SH-1).
MOD_RES 696 696 ALKYLATION (SH-2).
MOD_RES 706 706 ALKYLATION (SH-2).
CONFLICT 1331 1331 A -> G (IN REF. 3).
CONFLICT 1391 1392 KK -> QE (IN REF. 1 AND 2).
CONFLICT 1608 1609 SR -> RA (IN REF. 3).
CONFLICT 1663 1664 RG -> QT (IN REF. 2).
SEQUENCE 1940 AA; 224035 MW; 43CA58C6A4BA1253 CRC64;

Query Match 33.6%; Score 47; DB 1; Length 1940;
Best Local Similarity 47.1%; Pred. No. 36;

```

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 8 RVTHPHLPRALMRSTTK 24
| | | | | : : : : |
Db 664 RTTHPHEVRCIIPNETK 680

RESULT 12
MYH3_RAT
ID MYH3_RAT STANDARD; PRT; 1940 AA.
AC P12847;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOSIN HEAVY CHAIN, FAST SKELETAL MUSCLE, EMBRYONIC.
GN MYH3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87060988; PubMed=3783701;
RA Strehler E.E., Strehler-Page M.-A., Perriard J.C., Periasamy M.,
RA Nadal-Ginard B.;
RT "Complete nucleotide and encoded amino acid sequence of a mammalian
RT myosin heavy chain gene. Evidence against intron-dependent evolution
RT of the rod";
RL J. Mol. Biol. 190:291-317(1986).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X04267; CAA27817.1; -;
CC PIR; A24922; A24922.
CC HSSP; P13538; 2MYS.
CC InterPro: IPR000048; IQ.
CC InterPro: IPR002928; Myosin_tail.
CC InterPro: IPR002017; Spectrin.
CC InterPro: IPR001609; myosin_head.
CC Pfam; PF00612; IQ; 1.
CC Pfam; PF00063; myosin_head; 1.
CC Pfam; PF01576; Myosin_tail; 1.
CC PRINTS; PRO0193; MYOSINHEAVY.
CC ProDom; PD000355; myosin_head; 1.
CC SMART; SM00015; IQ; 1.
CC SMART; SM00242; MYSC; 1.
CC PROSITE; PS00096; IQ; 1.
CC Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
CC Calmodulin-binding; ATP-binding; Methylation; Alkylation;
CC Multigene family.
CC 1 781 MYOSIN HEAD-LIKE.
CC DOMAIN

FT DOMAIN 782 811 IQ.
FT DOMAIN 840 1933 COILED COIL (POTENTIAL).
FT NP_BIND 179 186 ATP (POTENTIAL).
FT DOMAIN 656 678 ACTIN-BINDING.
FT DOMAIN 758 772 ACTIN-BINDING.
FT MOD_RES 130 130 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 696 696 ALKYLATION (SH-1).
FT MOD_RES 706 706 ALKYLATION (SH-2).
SQ SEQUENCE 1940 AA; 223857 MW; B5D546A596ESA696 CRC64;

Query Match 33.6%; Score 47; DB 1; Length 1940;
Best Local Similarity 47.1%; Pred. No. 36;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 8 RVTHPHLPRALMRSTTK 24
| | | | | : : : : |
Db 664 RTTHPHEVRCIIPNETK 680

RESULT 13
IRS1_MOUSE
ID IRS1_MOUSE STANDARD; PRT; 1233 AA.
AC P35569;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE INSULIN RECEPTOR SUBSTRATE-1.
GN IRS1 OR IRS-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94220494; PubMed=8167159;
RA Araki E., Haag B.L. III, Kahn C.R.;
RT "Cloning of the mouse insulin receptor substrate-1 (IRS-1) gene and
RT complete sequence of mouse IRS-1";
RL Blochim. Biophys. Acta 1221:353-356(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93192326; PubMed=8448209;
RA Keller S.R., Aebersold R., Garner C.W., Lienhard G.E.;
RT "The insulin-elicited 160 kDa phosphotyrosine protein in mouse
RT adipocytes is an insulin receptor substrate 1: identification by
RT cloning";
RL Blochim. Biophys. Acta 1172:323-326(1993).
CC -1- FUNCTION: MAY MEDIATE THE CONTROL OF VARIOUS CELLULAR PROCESSES
CC BY INSULIN. WHEN PHOSPHORYLATED BY THE INSULIN RECEPTOR BINDS
CC SPECIFICALLY TO VARIOUS CELLULAR PROTEINS CONTAINING SH2 DOMAINS
CC SUCH AS PHOSPHATIDYLINOSITOL 3-KINASE P85 SUBUNIT OR GRB-2.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PTB DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L24563; AAA39335.1; -;
CC EMBL; X69722; CAA49378.1; -;
CC PIR; S43514; S43514.
CC HSSP; P35568; IIRS.
CC MGD; MGI:99454; Irs1.
CC InterPro: IPR002404; Insulin_Recep_S-1.
CC InterPro: IPR001849; PH.
CC Pfam; PF02174; IRS; 1.
CC Pfam; PF00169; PH; 1.
CC PRINTS; PRO0628; INSULINRSI.

DR SMART; SM00233; PH; 1.
DR SMART; SM00310; PTBI; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Phosphorylation.
FT DOMAIN 12 115 PH.
FT DOMAIN 152 262 PTB.
FT DOMAIN 152 262 POLY-SER.
FT DOMAIN 875 680 POLY-GLN.
FT DOMAIN 872 877 POLY-GLY.
FT DOMAIN 1119 1128 POLY-PRO.
FT DOMAIN 1194 1198 PHOSPHORYLATION (BY CK2)
FT MOD_RES 99 99 PHOSPHORYLATION (BY CK2)
FT MOD_RES 460 460 PHOSPHORYLATION (BY INSR)
FT MOD_RES 608 608 PHOSPHORYLATION (BY INSR)
FT MOD_RES 628 628 PHOSPHORYLATION (BY INSR)
FT MOD_RES 891 891 PHOSPHORYLATION (BY INSR)
FT MOD_RES 935 935 PHOSPHORYLATION (BY INSR)
FT MOD_RES 983 983 PHOSPHORYLATION (BY INSR)
FT MOD_RES 1173 1173 PHOSPHORYLATION (BY INSR)
FT MOD_RES 1220 1220 PHOSPHORYLATION (BY INSR)
FT MOD_RES 1038 1039 MISSING (IN REF. 2).
FT MOD_RES 1182 1182 H -> R (IN REF. 2).
SQ SEQUENCE 1233 AA; 130723 MW; C0E9B2B9D90DADD87 CRC64;

Query Match 33.2%; Score 46.5; DB 1; Length 1233;
Best Local Similarity 43.5%; Pred. No. 26;
Matches 10; Conservative 2; Mismatches 4; Indels 7; Gaps 1;

QY 1 CGEYQSRVTH-----PHLPR 16
II : I : II
DB 816 CGARPSLTPHHVLPQLPR 838

RESULT 14
YK14_YEAST
ID YK14_YEAST STANDARD; PRT; 116 AA.
AC P36078;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE HYPOTHETICAL 13.6 KDA PROTEIN IN MDH1-VMA5 INTERGENIC REGION.
GN YKL084W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Pohl T.M., Pohl F.M.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; Z28084; CAA81922.1; -
DR PIR; S37909; S37909.
DR SGD; S0001567; YKL084W.
DR Hypothetical protein.
SQ SEQUENCE 116 AA; 13641 MW; 57464E5FD9A591F8 CRC64;

Query Match 32.9%; Score 46; DB 1; Length 116;
Best Local Similarity 45.0%; Pred. No. 2.6;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 6 QSRVTHPLRALMRSTTKC 25
III I III : :
DB 14 QSRVHWHLPKDVIAIRFKC 33

RESULT 15
TRY3_RAT
ID TRY3_RAT STANDARD; PRT; 247 AA.
AC P08426;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TRYPSIN III, CATIONIC PRECURSOR (EC 3.4.21.4) (PRETRYPSINOGEN III).
GN TRY3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87271609; PubMed=3607011;
RA Fletcher T.S., Alhadeff M., Craik C.S., Largman C.;
RT "Isolation and characterization of a cDNA encoding rat cationic
RT trypsinogen.";
RL Biochemistry 26:3081-3086(1987).
CC -I- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC -I- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL; M16624; AAA41985.1; -
DR PIR; A27547; A27547.
DR HSSP; P00763; IDPO.
DR MEROPS; S01.151; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Trypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW Multigene family.
FT SIGNAL 1 15
FT PROPEP 16 24 ACTIVATION PEPTIDE.
FT CHAIN 25 247 TRYPSIN III, CATIONIC.
FT ACT_SITE 64 64 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 108 108 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 201 201 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 31 161 BY SIMILARITY.
FT DISULFID 49 65 BY SIMILARITY.
FT DISULFID 133 234 BY SIMILARITY.
FT DISULFID 140 207 BY SIMILARITY.
FT DISULFID 172 186 BY SIMILARITY.
FT DISULFID 197 221 BY SIMILARITY.
FT SITE 195 195 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ SEQUENCE 247 AA; 26269 MW; D74892BAA584E4A8 CRC64;

32.98; Score 46; DB 1; Length 247;

4 TYQSRVTHPLPRALMRSTTKC 25

Search completed: March 4, 2002, 13:08:41

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:56:38 ; Search time 79.15 Seconds
(without alignments)
46.201 Million cell updates/sec

Title: US-09-701-623C-5

Perfect score: 140
Sequence: 1 CGETYQSRVTHPLRALMRSTTKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_17.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-----------|--------------------|
| 1 | 54 | 38.6 | 684 | 13 Q90544 | Q90544 ginglymosto |
| 2 | 53.5 | 38.2 | 333 | 10 Q92T61 | Q92T61 sambucus ni |
| 3 | 53.5 | 38.2 | 333 | 10 Q9SV55 | Q9SV55 sambucus ni |
| 4 | 53.5 | 38.2 | 333 | 10 Q9YS4 | Q9YS4 sambucus ni |
| 5 | 52 | 37.1 | 320 | 10 Q9LQ87 | Q9LQ87 arabidopsis |
| 6 | 52 | 37.1 | 333 | 2 Q9RUY0 | Q9RUY0 streptomyce |
| 7 | 50.5 | 36.1 | 333 | 10 Q92T60 | Q92T60 sambucus ni |
| 8 | 50 | 35.7 | 1107 | 12 Q88603 | Q88603 tobacco mos |
| 9 | 50 | 35.7 | 1601 | 12 Q88604 | Q88604 tobacco mos |
| 10 | 49 | 35.0 | 273 | 5 Q9VFL6 | Q9VFL6 drosophila |
| 11 | 49 | 35.0 | 277 | 10 Q9LTK6 | Q9LTK6 arabidopsis |
| 12 | 49 | 35.0 | 360 | 5 Q9V9X5 | Q9V9X5 drosophila |
| 13 | 49 | 35.0 | 840 | 13 Q918N3 | Q918N3 rana pipien |
| 14 | 48 | 34.3 | 544 | 11 Q60764 | Q60764 mus musculu |
| 15 | 48 | 34.3 | 1065 | 10 Q48948 | Q48948 arabidopsis |
| 16 | 48 | 34.3 | 1065 | 10 Q9PHK6 | Q9PHK6 arabidopsis |
| 17 | 48 | 34.3 | 2254 | 10 Q9LN02 | Q9LN02 arabidopsis |
| 18 | 47 | 33.6 | 242 | 12 Q39842 | Q39842 influenza c |
| 19 | 47 | 33.6 | 242 | 12 Q89465 | Q89465 influenza c |

| | | | | | |
|----|------|------|------|-----------|--------------------|
| 20 | 47 | 33.6 | 242 | 12 Q67390 | Q67390 influenza c |
| 21 | 47 | 33.6 | 242 | 12 Q9WAE1 | Q9WAE1 influenza c |
| 22 | 47 | 33.6 | 242 | 12 Q9WAD3 | Q9WAD3 influenza c |
| 23 | 47 | 33.6 | 242 | 12 Q9IQ57 | Q9IQ57 influenza c |
| 24 | 47 | 33.6 | 242 | 12 Q9IQ55 | Q9IQ55 influenza c |
| 25 | 47 | 33.6 | 314 | 11 Q9ESH2 | Q9ESH2 rattus norv |
| 26 | 47 | 33.6 | 374 | 12 Q67388 | Q67388 influenza c |
| 27 | 47 | 33.6 | 374 | 12 Q67389 | Q67389 influenza c |
| 28 | 47 | 33.6 | 374 | 12 Q67391 | Q67391 influenza c |
| 29 | 47 | 33.6 | 374 | 12 Q67392 | Q67392 influenza c |
| 30 | 47 | 33.6 | 374 | 12 Q96792 | Q96792 influenza c |
| 31 | 47 | 33.6 | 374 | 12 Q9YMB6 | Q9YMB6 influenza c |
| 32 | 47 | 33.6 | 374 | 12 Q9YMB5 | Q9YMB5 influenza c |
| 33 | 47 | 33.6 | 374 | 12 Q9WAE4 | Q9WAE4 influenza c |
| 34 | 47 | 33.6 | 374 | 12 Q9WAE2 | Q9WAE2 influenza c |
| 35 | 47 | 33.6 | 374 | 12 Q9WAD9 | Q9WAD9 influenza c |
| 36 | 47 | 33.6 | 374 | 12 Q9WAD7 | Q9WAD7 influenza c |
| 37 | 47 | 33.6 | 374 | 12 Q9WAD5 | Q9WAD5 influenza c |
| 38 | 47 | 33.6 | 374 | 12 Q9W8X9 | Q9W8X9 influenza c |
| 39 | 47 | 33.6 | 374 | 12 Q9W841 | Q9W841 influenza c |
| 40 | 47 | 33.6 | 840 | 13 Q918N4 | Q918N4 rana pipien |
| 41 | 47 | 33.6 | 841 | 13 Q918N5 | Q918N5 rana pipien |
| 42 | 47 | 33.6 | 1938 | 6 Q9GJP9 | Q9GJP9 oryctolagus |
| 43 | 47 | 33.6 | 1938 | 13 Q91BD7 | Q91BD7 seriola dum |
| 44 | 46.5 | 33.2 | 209 | 2 Q9Z317 | Q9Z317 bradyrhizob |
| 45 | 46 | 32.9 | 249 | 11 Q9CV69 | Q9CV69 mus musculu |

ALIGNMENTS

RESULT 1

Q90544 ID Q90544 PRELIMINARY; PRT; 684 AA.
AC Q90544;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE NOVEL ANTIGEN RECEPTOR PRECURSOR.
OS Ginglymostoma cirratum (Nurse shark).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidae; Orectolobiformes;
OC Ginglymostomatidae; Ginglymostoma.
OC NCBI_TaxID=7801;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RX MEDLINE=95183140; PubMed=7877689;
RA Greenberg A.S., Avila D., Hughes A., McKinney E.C.,
RA Flajnik M.F.;
RT "A new antigen receptor gene family that undergoes rearrangement and
RT extensive somatic diversification in sharks.";
RL Nature 374:168-173(1995).
CC - SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
DOMAIN.
CC EMBL; U18701; AAB48195.1; .
DR HSP; P01857; IFC1.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 6.
DR SMART; SM00407; IGcl; 4.
DR SMART; SM00410; IG_Like; 2.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Signal.
FT SIGNAL
FT CHAIN
SQ SEQUENCE 684 AA; 75224 MW; 2FF9D2071CDA6DFD CRC64;

Query Match 38.6%; Score 54; DB 13; Length 684;
Best Local Similarity 57.1%; Pred. No. 5;
Matches 12; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 4 TYQSRVTHPHLPRALMRSTTK 24
DB 642 TYSCLVGHPSLNRDLIRSTNK 662

RESULT 2

| | | | |
|--------|---------------------------------------------------------------------|------|---------|
| Q9ZT61 | PRELIMINARY; | PRT; | 333 AA. |
| ID | Q9ZT61 | | |
| AC | Q9ZT61; | | |
| DT | 01-MAY-1999 (TrEMBLrel. 10, Created) | | |
| DT | 01-MAY-1999 (TrEMBLrel. 10, Last sequence update) | | |
| DT | 01-JUN-2001 (TrEMBLrel. 17, Last annotation update) | | |
| DE | HEVEIN-LIKE PROTEIN. | | |
| DE | Sambucus nigra (European elder). | | |
| OS | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |
| OC | EC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | |
| OC | Asteridae; euasterids II; Adoxaceae; Sambucus. | | |
| OX | NCBI_TaxID=4202; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | TISSUE=ELDERBERRY FRUIT; | | |
| RA | van Damme E.J.M., Charrels D., Roy S., Tierens K., Barre A., | | |
| RA | Martins J.C., Rouge P., Van Leuven F., Does M., Peumans W.J.; | | |
| RT | "A gene consisting of the N-terminal domain of a PR-4 protein and | | |
| RT | C-terminal domain of a class V chitinase PR-3 encodes a hevein-like | | |
| RT | protein in elderberry (Sambucus nigra L.) fruits."; | | |
| RL | Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases. | | |
| RL | EMBL; AF074385; AAD11406.1; ; | | |
| DR | HSP: P02877; 1HEV. | | |
| DR | InterPro: IPR001002; Chitin_bind | | |
| DR | InterPro: IPR000726; Glyco_hydro_19.. | | |
| DR | Pfam: PF00187; Chitin_Binding; 1. | | |
| DR | Pfam: PF00182; Glyco_hydro_19; 1. | | |
| DR | ProDom: PD003574; Glyco_hydro_19; 1. | | |
| DR | ProDom: PD000609; Chitin_bind; 1. | | |
| DR | SMART: SM00270; ChcBD1; 1 | | |
| DR | PROSITE: PS00774; CHITINASE_19.2; 1. | | |
| DR | PROSITE: PS00026; CHITIN_BINDING; 1. | | |
| DR | Chitin-binding. | | |
| KW | Chitin-binding. | | |
| SC | SEQUENCE 333 AA; 37077 MW; 3231A252FC253BE5 CRC64; | | |

Query Match 38.2%; Score 53.5; DB 10; Length 333;
Best Local Similarity 42.9%; Pred. No. 3;
Matches 15; Conservative 2; Mismatches 3; Indels 15

```
Qy 1 CGETYQ-----SRVTHPLPRALMRST 22
    || |||
Db 52 CGSTYQKCEDGCQSQCRDTSRLT--DLPRALLRPT 84
```

RESULT

| Accession | Protein | Length | Score | E-value | Ident | Query | Accession | Protein | Length | Score | E-value | Ident | Query |
|-----------|--------------------------------------------------------------------|--------|-------|---------|-------|-------|-----------|--------------------------------------------------------------------|--------|-------|---------|-------|-------|
| Q9SY55 | PRELIMINARY; | | | | | | Q9SY55 | PRELIMINARY; | | | | | |
| ID | Q9SY55 | | | | | | ID | Q9SY55 | | | | | |
| AC | Q9SY55; | | | | | | AC | Q9SY55; | | | | | |
| DT | 01-MAY-2000 (TrEMBLrel. 13, Created) | | | | | | DT | 01-MAY-2000 (TrEMBLrel. 13, Created) | | | | | |
| DT | 01-MAY-2000 (TrEMBLrel. 13, Last sequence update) | | | | | | DT | 01-MAY-2000 (TrEMBLrel. 13, Last sequence update) | | | | | |
| DT | 01-JUN-2001; (TrEMBLrel. 17, Last annotation update) | | | | | | DT | 01-JUN-2001; (TrEMBLrel. 17, Last annotation update) | | | | | |
| DE | HEVEIN-LIKE PROTEIN HLPF (FRAGMENT). | | | | | | DE | HEVEIN-LIKE PROTEIN HLPF (FRAGMENT). | | | | | |
| OS | Sambucus nigra (European elder). | | | | | | OS | Sambucus nigra (European elder). | | | | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | | | | | OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | | | | |
| OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | | | | | OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | | | | |
| OC | Asteridae; euasterids II; Adoxaceae; Sambucus. | | | | | | OC | Asteridae; euasterids II; Adoxaceae; Sambucus. | | | | | |
| OX | NCBI TaxID:4202; | | | | | | OX | NCBI TaxID:4202; | | | | | |

AN
RP
[1]
SEQUENCE FROM N. A.

SEQUENCE FROM N.A.
RC TISSUE=FRUIT;
RX MEDLINE=95214486; PubMed=10198114;
RA Van Damme E.J.M., Charrels D., Roy S., Tiersens K., Barre A.,
RA Martins J.C., Ronge P., Van Leuven F., Does M., Peumans W.D.;
RT "A gene encoding a hevein-like protein from eiderberry fruits is
RT homologous to PR-4 and class V chitinase genes."

| | | | | |
|----|-----------------|------------|-------------------|--------|
| RL | Plant Physiol. | 119: | 1547-1556(1999). | |
| DR | EMBL; | AF274388; | AAD12237.1; | - |
| DR | HSSB; | P02877; | 1HEV. | |
| DR | InterPro; | IPR001002; | Chitin_bind. | |
| DR | InterPro; | IPR000726; | Glyco_hydro_19. | |
| DR | Pfam; | PF00187; | chitin_binding_1. | |
| DR | Pfam; | PF00182; | Glyco_hydro_19_1. | |
| DR | PRINTS; | PR00451; | CHITINBIND. | |
| DR | ProDom; | PD000574; | Glyco_hydro_19_1. | |
| DR | ProDom; | PD006069; | Chitin_bind_1. | |
| DR | SMART; | SM00270; | ChitBD1_1. | |
| DR | PROSITE; | PS00774; | CHITINASE_19_2_1. | |
| DR | PROSITE; | PS00026; | CHITIN_BINDING_1. | |
| KW | Chitin-binding. | | | |
| FT | NON_TER | 333 | 333 | |
| SQ | SEQUENCE | 333 AA; | 37047 MW; | 391322 |

| | | | | |
|--------------------------|--------|---------------|------------|-------------|
| Query Match | 38.2% | Score 53.5; | DB 10; | Length 333; |
| Best Local Similarity | 42.9%; | Pred. No. 3; | | |
| Matches 15; Conservative | 2; | Mismatches 3; | Indels 15; | Gaps 2; |

```
Qy 1 CGETYQ-----SRVTHPLRALMRST 22
    || |||
Db 52 CGSTYQYCEGCGSQCRDTSRLT--DLPRALLRPT 84
```

RESULT 4

| | | | | | |
|------------------|--------------------------------------------------------------------|----------------------------------------|---------|--|--|
| Q9YSY4 | | PRT; | 333 AA. | | |
| ID | Q9SYS4 | PRELIMINARY; | | | |
| AC | Q9SY54 | | | | |
| DT | 01-MAY-2000 | (TREMBRel. 13, Created) | | | |
| DT | 01-MAY-2000 | (TREMBRel. 13, Last sequence update) | | | |
| DT | 01-JUN-2001 | (TREMBRel. 17, Last annotation update) | | | |
| DE | HEVEIN-LIKE PROTEIN. | | | | |
| OS | <i>Sambucus nigra</i> (European elder). | | | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | | | |
| OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | | | |
| OX | Asteridae; euasterids II; Adoxaceae; Sambucaceae; Sambucus. | | | | |
| NCBI_TaxID=4202; | | | | | |
| {1} | {1} | | | | |
| RN | SEQUENCE FROM N.A. | | | | |
| RP | | | | | |

TX TISSUE=ELDERBERRY FRUIT;
 RC MEDLINE=99214486; PubMed=10198114;
 RA Van Damme E.J.M., Charlels D., Roy S., Tierens K., Barre A.,
 RA Martins J.C., Rouge P., Van Leuven F., Does M., Peumans W.J.;
 RA "A gene encoding a hevein-like protein from elderberry fruits is
 RT homologous to PR-4 and class V chitinase genes.";
 RL Plant Physiol. 119:1547-1556(1999).

DR HSP; P02877; IHEV.
DR EMBL; AF074387; RAD11408.1;
DR InterPro; IPR001002; Chitin_bind.
DR InterPro; IPR000726; Glyco_hydro_19.
DR Pfam; PF00187; chitin_binding; 1.
DR Pfam; PF00182; Glyco_hydro_19; 1.
DR ProDom; PD000574; Glyco_hydro_19; 1.
DR ProDom; PD000609; Chitin_bind; 1.
DR SMART; SM00270; ChbSD1; 1.
DR PROSITE; PS00774; CHITINASE_19.2; 1.
DR PROSITE; PS00026; CHITIN_BINDING; 1.
KW Chitin-binding.
SQ SEQUENCE 333 AA; 37126 MW; 3231B1

Query Match 38.2%; Score 53.5; DB 10; Length 333;
Best Local Similarity 42.9%; Pred. No. 3;
Matches 15: Conservative 2; Mismatches 3; Indels 15; Gaps 2;

QY 1 CGETYQ-----SRVTHPLPALMRST 22
 || ||| ||:| ||||:| |
 Db 52 CGSTYQVCEGCGSQCRDTSRLT--DLPRALLRPT 84

```
RESULT 5
Q9LQ87 ID Q9LQ87 PRELIMINARY; PRT; 320 AA.
AC Q9LQ87
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE T1N6.10 PROTEIN.
DE T1N6.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Liu S., Chan A., Yu G., Lee J., Lenz C., Pham P., Sakano H.,
RA Toriumi M., Chin C., Chlou J., Choi E., Chung M., Gonzalez A.,
RA Howing B., Liu A., Vaysberg M., Altafi H., Brooks S., Buehler E.,
RA Chao Q., Conn L., Conway A., Hansen N., Johnson-Hopson C., Khan S.,
RA Kim C., Lam B., Miranda M., Nguyen M., Palm C., Shinn P.,
RA Southwick A., Davis R., Ecker J., Federspiel N., Theologis A.;
RT "The sequence of BAC T1N6 from Arabidopsis thaliana chromosome 1.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC009273; AAF78401.1; -
DR InterPro; IPR003703; Acyl_CoA_thio.
DR Pfam; PF02551; Acyl_CoA_thio; 1.
SQ SEQUENCE 320 AA; 36157 MW; 03A82A96333E1BFC CRC64;

Query Match 37.1%; Score 52; DB 10; Length 320;
Best Local Similarity 60.0%; Pred. No. 4.9;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 ETYQSRVTHPLPRA 17
| :||:| |||||
Db 142 ELRESRTDPLPRS 156

RESULT 6
Q9RJY0 ID Q9RJY0 PRELIMINARY; PRT; 533 AA.
AC Q9RJY0
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PUTATIVE EXPORT PROTEIN.
DE SCF37.21C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL133210; CAB61602.1; -
```

```
DR InterPro; IPR001411; TCR_TetB.
DR PRINTS; PRO1036; TCRTEB.
SQ SEQUENCE 533 AA; 54474 MW; 49C650E9797FEF77 CRC64;

Query Match 37.1%; Score 52; DB 2; Length 533;
Best Local Similarity 58.8%; Pred. No. 7.9;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 5 YQSRVTHPLPRLMRS 21
| :||:| |||||
Db 268 YETRVHPLPMLFRS 284

RESULT 7
Q9ZT60 ID Q9ZT60 PRELIMINARY; PRT; 333 AA.
AC Q9ZT60
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HEVEIN-LIKE PROTEIN.
OS Sambucus nigra (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Adoxaceae; Sambucus.
OX NCBI_TaxID=4202;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ELDERBERRY FRUIT;
RA van Damme E.J.M., Charels D., Roy S., Tierens K., Barre A.,
RA Martins J.C., Rouge P., Van Leuven F., Does M., Peumans W.J.;
RT "A gene consisting of the N-terminal domain of a PR-4 protein and the
RT C-terminal domain of a class V chitinase PR-3 encodes a hevein-like
RT protein in elderberry (Sambucus nigra L.) fruits.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF074386; AAD11407.1; -
DR HSSP; P02877; IHEV.
DR InterPro; IPR001002; Chitin_bind.
DR InterPro; IPR000726; Glyco_hydro_19.
DR Pfam; PF00187; chitin_binding; 1.
DR Pfam; PF00182; Glyco_hydro_19; 1.
DR ProDom; PD000574; Glyco_hydro_19; 1.
DR ProDom; PD000609; Chitin_bind; 1.
DR SMART; SM00270; ChtBD1; 1.
DR PROSITE; PS00774; CHITINASE_19_2; 1.
DR PROSITE; PS00026; CHITIN_BINDING; 1.
KW Chitin-binding.
SQ SEQUENCE 333 AA; 36922 MW; 70C6969392A1174A CRC64;

Query Match 36.1%; Score 50.5; DB 10; Length 333;
Best Local Similarity 40.0%; Pred. No. 8.5;
Matches 14; Conservative 3; Mismatches 3; Indels 15; Gaps 2;

QY 1 CGETYQ-----SRVTHPLPRLMRST 22
| :||:| |||||
Db 52 CGSTYQYCEDGCQCQCRDARLT--DLPRALLRPT 84

RESULT 8
Q88603 ID Q88603 PRELIMINARY; PRT; 1107 AA.
AC Q88603
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE DE 122K.
OS Tobacco mosaic virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
OX NCBI_TaxID=12242;
RN [1]
RP SEQUENCE FROM N.A.
```



```

QY 11 HPHLPALMRSTTKC 25
Db 100 HPHLPALPANLQRC 114

RESULT 11
Q9LTK6 PRELIMINARY; PRT; 277 AA.
AC Q9LTK6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE GENOMIC DNA, CHROMOSOME 5, BAC CLONE:FL7P19.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eusoids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RA Kaneko T., Kato T., Asamizu E., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB025603; BAA97457.1; -.
SQ SEQUENCE 277 AA; 31592 MW; C72F0D844D6A2089 CRC64;

Query Match 35.0%; Score 49; DB 10; Length 277;
Best Local Similarity 36.4%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 2 GERYQSRVTHPLPALMRSTT 23
Db 148 GENESGRTYTHSPSSISHNPT 169

RESULT 12
Q9V9X5 PRELIMINARY; PRT; 360 AA.
AC Q9V9X5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE CG11335 PROTEIN (LYSYL OXIDASE-LIKE).
GN LOX OR CG11335.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthanathan P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.J., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram J.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busan D.A., Butler H., Cadenot L.B., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong J., Gorrell J.H., Gu Z., Guan P., Harris M.,

```

```

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Smith H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Mink M., Csizsar K.;
RT "The role of Drosophila lysyl oxidase in the regulation of chromatin
RT structure.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003778; AAF57153.1; -.
DR EMBL; AJ295625; CAB99480.1; -.
DR FlyBase; FBgn0039848; lox.
DR InterPro; IPR001695; Lysyl_oxidase.
DR InterPro; IPR001190; SRCR.
DR Pfam; PF01186; Lysyl_oxidase; 1.
DR Pfam; PF00530; SRCR; 1.
DR PRINTS; PR00074; LYSYLOXIDASE.
DR PRINTS; PR00258; SPERACTRCPTR.
DR ProDom; PD013887; Lysyl_oxidase; 1.
DR SMART; SM00202; SR; 1.
DR PROSITE; PS0287; SRCR_2; 1.
SQ SEQUENCE 360 AA; 40832 MW; 7130F7265CB5B698 CRC64;

Query Match 35.0%; Score 49; DB 5; Length 360;
Best Local Similarity 40.9%; Pred. No. 15;
Matches 9; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 ETQSRVTHPLPALMRSTTK 24
Db 184 DAYEIRTNPHAAKILLRFSVK 205

RESULT 13
Q918N3 PRELIMINARY; PRT; 840 AA.
AC Q918N3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE TYPE 3 MYOSIN HEAVY CHAIN (FRAGMENT).
OS Rana pipiens (Northern leopard frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
OX NCBI_TaxID=8404;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ANTERIOR TIBIALIS SKELETAL MUSCLE;
RX MEDLINE=20314481; PubMed=10854783;
RA Lutz G.J., Razzaghi S., Lieber R.L.;
RT "Cloning and characterization of the S1 domain of four myosin isoforms
RT from functionally divergent fiber types in adult Rana pipiens skeletal
RT muscle.";
RL Gene 250:97-107(2000).
DR EMBL; AF240690; AAF82092.1; -.

```


THIS PAGE BLANK (us:17)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:58:20 ; Search time 42.32 Seconds
(without alignments)
44.999 Million cell updates/sec

Title: US-09-701-623C-6
Perfect score: 141
Sequence: 1 CGETYYSRVTTHPLPKDIVRSIAKC 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|------------------------------------------|
| 1 | 88 | 62.4 | 426 | 2 I36948 | Ig epsilon-chain - chimpanzee (fragment) |
| 2 | 85 | 60.3 | 428 | 1 EHHU | Ig epsilon chain C |
| 3 | 82 | 58.2 | 429 | 1 EHRT | Ig epsilon chain C3 regio |
| 4 | 58 | 41.1 | 107 | 2 I68730 | IgE chain C3 regio |
| 5 | 58 | 41.1 | 107 | 2 I68726 | Ig epsilon chain C |
| 6 | 58 | 41.1 | 388 | 1 EHMS | Ig epsilon chain C |
| 7 | 58 | 41.1 | 548 | 2 S38864 | Ig epsilon chain C |
| 8 | 54 | 38.3 | 116 | 2 S37909 | hypothetical prote |
| 9 | 54 | 38.3 | 394 | 2 T19571 | hypothetical prote |
| 10 | 54 | 38.3 | 1306 | 2 S42659 | acetyl-CoA carboxy |
| 11 | 54 | 38.3 | 1625 | 2 T02921 | acetyl-CoA carboxy |
| 12 | 54 | 38.3 | 2311 | 2 T06161 | acetyl-CoA carboxy |
| 13 | 54 | 38.3 | 2325 | 2 T02235 | acetyl-CoA carboxy |
| 14 | 52 | 36.9 | 1756 | 2 T02599 | hypothetical prote |
| 15 | 51.5 | 36.5 | 605 | 1 FPMS | alpha-fetoprotein |
| 16 | 51 | 36.2 | 106 | 2 A03859 | hypothetical prote |
| 17 | 51 | 36.2 | 195 | 2 A75171 | hypothetical prote |
| 18 | 50 | 35.5 | 539 | 2 T48197 | hypothetical prote |
| 19 | 49.5 | 35.1 | 381 | 2 S28115 | gas-vesicle protei |
| 20 | 49.5 | 35.1 | 382 | 2 T01122 | gas-vesicle protei |
| 21 | 49.5 | 35.1 | 382 | 2 T08243 | gas-vesicle operon |
| 22 | 49 | 34.8 | 684 | 2 S60266 | novel antigen rece |
| 23 | 49 | 34.8 | 1685 | 2 T02750 | acetyl-CoA carboxy |
| 24 | 48 | 34.0 | 384 | 2 T06814 | acetyl-CoA carboxy |
| 25 | 48 | 34.0 | 454 | 2 A46532 | Ig mu chain C regi |
| 26 | 48 | 34.0 | 958 | 2 S15566 | pol protein - simi |
| 27 | 48 | 34.0 | 1161 | 2 S18738 | Ig epsilon chain - |
| 28 | 47 | 33.3 | 110 | 2 S43147 | Ig epsilon chain - |
| 29 | 47 | 33.3 | 320 | 2 C86148 | hypothetical prote |

30 47 33.3 372 2 T16971 probable GTP-bindin
31 47 33.3 384 1 RGTOOA GTP-binding regula
32 47 33.3 2254 2 D86215 protein T6D22.14 f
33 47 33.3 2257 2 T09538 acetyl-CoA carboxy
34 46.5 33.0 248 2 T48376 hypothetical prote
35 46 32.6 198 2 B71044 hypothetical prote
36 46 32.6 245 2 I55951 MHC class II E-bet
37 46 32.6 264 2 A60497 H-2 class II histo
38 46 32.6 264 2 S10989 class II histocomp
39 46 32.6 382 2 T34940 probable membrane
40 46 32.6 549 2 S04845 Ig heavy chain pre
41 45.5 32.3 270 1 A64567 pyruvate synthase
42 45 31.9 242 1 MFIVCJ matrix protein M1
43 45 31.9 319 1 S57926 protein kinase (EC
44 45 31.9 365 2 T04718 hypothetical prote
45 45 31.9 372 2 D64716 regulatory protein

ALIGNMENTS

RESULT 1
I36948
Ig epsilon-chain - chimpanzee (fragment)
C:Species: Pan troglodytes (chimpanzee)
C>Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jan-2000
C:Accession: I36948
R:Sakoyama, Y.; Hong, K.
Proc. Natl. Acad. Sci. U.S.A. 84, 1080-1084, 1987
A:Title: Nucleotide sequences of immunoglobulin epsilon genes of chimpanzee and orang
A:Reference number: I36948; MUID:87147196
A:Accession: I36948
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-426 <RES>
A:Cross-references: GB:M15398; NID:gl76797; PIDN:AAA35416.1; PID:gl76798
C:Genetics:
A:Introns: 103/1; 209/1; 317/1
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:336-405/Domain: immunoglobulin homology <IMM>

Query Match 62.4%; Score 88; DB 2; Length 426;
Best Local Similarity 69.6%; Pred. No. 3.6e-05;
Matches 16; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 2 GETYYSRVTTHPLPKDIVRSIAK 24
||||| |||||||: ||| |
Db 292 GETYQCRVTHPLPRALVRSTTK 314
RESULT 2
EHHU
Ig epsilon chain C region - human
C:Species: Homo sapiens (man)
C>Date: 31-Mar-1981 #sequence_revision 13-Jun-1983 #text_change 16-Jul-1999
C:Accession: A22771; A23195; PH1214; A93491; A90824; A94418; B93933; S02438; A53116;
R:Flanagan, J.G.; Rabbitts, T.H.
EMBO J. 1, 655-660, 1982
A:Title: The sequence of a human immunoglobulin epsilon heavy chain constant region 9
A:Reference number: A22771; MUID:84236029
A:Accession: A22771
A:Molecule type: DNA
A:Residues: 1-428 <FLA>
A:Cross-references: GB:L00022; GB:J00227; GB:V00555; NID:gl85035
R:Ueda, S.; Nakai, S.; Nishida, Y.; Hisajima, H.; Honjo, T.
EMBO J. 1, 1539-1544, 1982
A:Title: Long terminal repeat-like elements flank a human immunoglobulin epsilon pseu
A:Reference number: A23195; MUID:84207910
A:Accession: A23195
A:Molecule type: DNA
A:Residues: 2-428 <UED>
A:Cross-references: GB:J00222; NID:gl84755

A:Cross-references: GB:S55276; NID:g263168; PIDN:AAB24858.1; PID:g263169
A:Experimental source: B cell myeloma U-266
A:Note: sequence extracted from NCBI backbone (NCBIP:125299)
A:Accession: A46536
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 401-428 <HE>
A:Cross-references: GB:S53497; NID:g263162; PIDN:AAB24855.1; PID:g263163
A:Experimental source: B cell myeloma U-266
A:Note: sequence extracted from NCBI backbone (NCBIP:123483)
C:Genetics:
A:Gene: GDB:IGHE
A:Cross-references: GDB:119335; OMIM:147180
A:Map position: 14q32.33-14q32.33
A:Introns: 1/1; 104/1; 211/1; 319/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobulin homology
F:22-87/Domain: immunoglobulin homology <IM1>
F:128-195/Domain: immunoglobulin homology <IM2>
F:232-301/Domain: immunoglobulin homology <IM3>
F:338-407/Domain: immunoglobulin homology <IM4>
F:14/Disulfide bonds: interchain (to light chain) #status predicted
F:15-105,29-85,135-193,239-299,345-405/disulfide bonds: #status predicted
F:21,49,99,146,252,275/Binding site: carbohydrate (Asn) (covalent) #status experiment
F:121,209/Disulfide bonds: interchain (to heavy chain) #status predicted

Query Match 60.3%; Score 85; DB 1; Length 428;
Best Local Similarity 65.2%; Pred. No. 9.9e-05;
Matches 15; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GETVYSRVTHPLPKDITVRSIAK 24
||||| ||||| :||| |

Db 294 GETYQCRVTHPLPRALMRSTTK 316

RESULT 3
EHRT
Ig epsilon chain C region - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 16-Jul-1999
C:Accession: A93442; A90937; A02143
R:Hellman, L.; Pettersson, U.; Engstrom, A.; Karlsson, T.; Bennich, H.
Nucleic Acids Res. 10, 6041-6049, 1982
A:Title: Structure and evolution of the heavy chain from rat immunoglobulin E.
A:Reference number: A93442; MUID:83064537
A:Accession: A93442
A:Molecule type: mRNA
A:Residues: 1-429 <HE>
A:Experimental source: strain LOU/c/Wsl, immunocytoma IR2
R:Kindsvogel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.
DNA 1, 335-343, 1982
A:Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction
A:Reference number: A90937; MUID:83182019
A:Contents: myeloma IR162
A:Accession: A90937
A:Molecule type: mRNA
A:Residues: 'N',169-307,'L',309-342 <KIN>
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:19-80/Domain: immunoglobulin homology <IM1>
F:118-186/Domain: immunoglobulin homology <IM2>
F:223-291/Domain: immunoglobulin homology <IM3>
F:327-398/Domain: immunoglobulin homology <IM4>
F:46,99,170,240,265,369,419/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 58.2%; Score 82; DB 1; Length 429;
Best Local Similarity 69.6%; Pred. No. 0.00027;

Matches 16; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIVRSIAK 24

I I I I I I I I I I I I I I I I

Db 284 GEGYQCRVDHPFPKPIVRSITK 306

RESULT 4

IgE chain C3 region - mouse (fragment)

A:Species: Mus musculus (house mouse)

C:Date: 02-Aug-1996 #sequence_revision

C:Accession: I68730 #text_change 21-Jan-2000

R:Shinkai, Y.; Nakauchi, H.; Honjo, T.; Okumura, K.

Immunogenetics 27, 288-292, 1988

A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic acid

A:Reference number: I54443; MUID:88152907

A:Accession: I68730

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-107 <RES>

A:Cross-references: GB:M22933; NID:g194464; PIDN:AAA37915.1; PID:g194469

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F:22-90/Domain: immunoglobulin homology <IMM>

Query Match 41.1%; Score 58; DB 2; Length 107;

Best Local Similarity 56.5%; Pred. No. 0.21;

Matches 13; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIVRSIAK 24

I I I I I I I I I I I I I I I I

Db 83 GYGQCIVDHPFPKPIVRSITK 105

RESULT 5

I68726

IgE chain C3 region - mouse (fragment)

A:Species: Mus musculus (house mouse)

C:Date: 02-Aug-1996 #sequence_revision

C:Accession: I68726

R:Shinkai, Y.; Nakauchi, H.; Honjo, T.; Okumura, K.

Immunogenetics 27, 288-292, 1988

A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic acid

A:Reference number: I54443; MUID:88152907

A:Accession: I68726

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-107 <RES>

A:Cross-references: GB:M22930; NID:g194455; PIDN:AAA37911.1; PID:g194460

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F:22-90/Domain: immunoglobulin homology <IMM>

Query Match 41.1%; Score 58; DB 2; Length 107;

Best Local Similarity 56.5%; Pred. No. 0.21;

Matches 13; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIVRSIAK 24

I I I I I I I I I I I I I I I I

Db 83 GYGQCIVDHPFPKPIVRSITK 105

RESULT 6

EHMS

Ig epsilon chain C region (version 1) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 13-Jun-1983 #sequence_revision

C:Accession: A02144 #text_change 16-Jul-1999

R:Li, F.T.; Albrandt, K.; Sutcliffe, J.G.; Katz, D.H.

Proc. Natl. Acad. Sci. U.S.A. 79, 7852-7856, 1982

A:Title: Cloning and nucleotide sequence of mouse immunoglobulin epsilon chain cDNA.

A:Reference number: A02144; MUID:83117774

Query Match

Best Local Similarity

38.3%; Score 54; DB 2; Length 116;

Pred. No. 0.87;

A:Accession: A02144

A:Molecule type: mRNA

A:Residues: 1-388 <LIU>

A:Cross-references: GB:J00476; NID:g194875; PIDN:AAA38085.1; PID:g387220

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (

hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotetramer; heterotetramer; immunoglobulin

F:1-44/Domain: immunoglobulin homology (fragment) <IM1>

F:81-149/Domain: immunoglobulin homology <IM2>

F:186-254/Domain: immunoglobulin homology <IM3>

F:290-361/Domain: immunoglobulin homology <IM4>

F:10,51,62,133,205,228,332,382/Binding site: carbohydrate (Asn) (covalent) #status pr

Query Match

Best Local Similarity

41.1%; Score 58; DB 1; Length 388;

Matches 13; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIVRSIAK 24

I I I I I I I I I I I I I I I I

Db 247 GYGQCIVDHPFPKPIVRSITK 269

RESULT 7

S38864

Ig epsilon chain C region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 03-Nov-2000

C:Accession: S38864

R:Klipp, B.; Becker, W.; Schlaak, M.

submitted to the EMBL Data Library, November 1993

A:Description: Combination of a defined specificity and desired isotype by cloning of

A:Reference number: S38864

A:Accession: S38864

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-548 <KIP>

A:Cross-references: EMBL:Z27397; NID:g416537; PIDN:CAA81788.1; PID:g940782

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F:353-421/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity

41.1%; Score 58; DB 2; Length 548;

Matches 13; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GETYYSRVTHPLPKDIVRSIAK 24

I I I I I I I I I I I I I I I I

Db 414 GYGQCIVDHPFPKPIVRSITK 436

RESULT 8

S37909

hypothetical protein YKL084w - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C:Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 29-Oct-1999

C:Accession: S37909

R:Pohl, T.M.; Pohl, F.M.

submitted to the Protein Sequence Database, March 1994

A:Reference number: S37897

A:Accession: S37909

A:Molecule type: DNA

A:Residues: 1-116 <POH>

A:Cross-references: EMBL:Z28084; NID:g486122; PID:g486123; GSPDB:GN00011; MIPS:YKL084

A:Experimental source: strain S288C

C:Genetics:

A:Gene: MIPS:YKL084w

A:Map position: 11L

Query Match

Best Local Similarity

38.3%; Score 54; DB 2; Length 116;

Pred. No. 0.87;

C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C;Accession: T02921

Search completed: March 4, 2002, 12:58:20
Job time: 247 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:08:41 ; Search time 24.65 Seconds
(without alignments)
37.185 Million cell updates/sec

Title: US-09-701-623C-6
Perfect score: 141
Sequence: 1 CGETYYSRVTHPLPKDIVRSIAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36564827 residues
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|---------------|---------------------|
| 1 | 85 | 60.3 | 428 | 1 EPC_HUMAN | P01854 homo sapien |
| 2 | 82 | 58.2 | 429 | 1 EPC_RAT | P01855 rattus norv |
| 3 | 58 | 41.1 | 421 | 1 EPC_MOUSE | P06336 mus musculus |
| 4 | 54 | 38.3 | 116 | 1 YK14_YEAST | P36078 saccharomyc |
| 5 | 51.5 | 36.5 | 605 | 1 FETA_MOUSE | P02772 mus musculus |
| 6 | 51 | 36.2 | 106 | 1 Y116_ADE02 | P03287 human adeno |
| 7 | 49.5 | 35.1 | 381 | 1 GVPC_HALME | P02228 halobacteri |
| 8 | 49.5 | 35.1 | 382 | 1 GVC1_HALN1 | P24574 halobacteri |
| 9 | 48 | 34.0 | 384 | 1 GBA2_PEA | P04279 pisum sativ |
| 10 | 48 | 34.0 | 1161 | 1 POL_SFV1 | P23074 simian foam |
| 11 | 47 | 33.3 | 384 | 1 GBA1_LYCES | P26981 lycopersico |
| 12 | 47 | 33.3 | 777 | 1 BAR1_HUMAN | P09728 homo sapien |
| 13 | 46 | 32.6 | 264 | 1 HB2D_RAT | P18211 rattus norv |
| 14 | 46 | 32.6 | 607 | 1 UVRC_PSEFL | P32966 pseudomonas |
| 15 | 45.5 | 32.3 | 270 | 1 KORB_METJA | Q57957 methanococc |
| 16 | 45 | 31.9 | 242 | 1 VMAT_INCIJJ | P12446 influenza c |
| 17 | 45 | 31.9 | 319 | 1 CDK4_XENLA | Q91727 xenopus lae |
| 18 | 45 | 31.9 | 626 | 1 UVRC_RICPR | Q92CX9 rickettsia |
| 19 | 44.5 | 31.6 | 1235 | 1 IRS1_RAT | P35570 rattus norv |
| 20 | 44 | 31.2 | 285 | 1 GVC2_HALN1 | Q9HHT0 halobacteri |
| 21 | 44 | 31.2 | 285 | 1 GVC2_HALN2 | P48310 halobacteri |
| 22 | 44 | 31.2 | 374 | 1 SPOP_HUMAN | O43791 homo sapien |
| 23 | 43.5 | 30.9 | 450 | 1 MUC_CANFA | P01874 canis famil |
| 24 | 43.5 | 30.9 | 530 | 1 CG2B_DROME | P20439 drosophila |
| 25 | 43 | 30.5 | 130 | 1 ALB1_PEA | P08687 pisum sativ |
| 26 | 43 | 30.5 | 260 | 1 CAH1_MOUSE | P13634 mus musculus |
| 27 | 43 | 30.5 | 275 | 1 TFS0_METJA | Q58170 methanococc |
| 28 | 43 | 30.5 | 376 | 1 YP60_MOUSE | Q88856 mus musculus |
| 29 | 43 | 30.5 | 391 | 1 MUCB_HUMAN | P04220 homo sapien |
| 30 | 43 | 30.5 | 418 | 1 YW13_YEAST | Q04526 saccharomyc |
| 31 | 43 | 30.5 | 454 | 1 MUC_HUMAN | P01871 homo sapien |
| 32 | 43 | 30.5 | 515 | 1 G6PD_TREPA | O83491 treponema p |
| 33 | 43 | 30.5 | 1095 | 1 PIPA_DROME | P13217 drosophila |

| RESULT | 1 | ALIGNMENTS |
|-----------|------------------------------------------------------------------------|------------------------|
| EPC_HUMAN | | |
| ID | EPC_HUMAN | STANDARD; PRT; 428 AA. |
| AC | P01854; | |
| DT | 21-JUL-1986 (Rel. 01, Created) | |
| DT | 21-JUL-1986 (Rel. 01, Last sequence update) | |
| DT | 20-AUG-2001 (Rel. 40, Last annotation update) | |
| DE | IG EPSILON CHAIN C REGION. | |
| GN | IGHE. | |
| OS | Homo sapiens (Human). | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | |
| OX | NCBI_TaxID=9606; | |
| RN | [1] | |
| RP | SEQUENCE FROM N.A. | |
| RX | MEDLINE=83168897; PubMed=6300763; | |
| RA | Seno M., Kurokawa T., Ono Y., Onda H., Sasada R., Igarashi K., | |
| RA | Kikuchi M., Sugino Y., Nishida Y., Honjo T.; | |
| RT | "Molecular cloning and nucleotide sequencing of human immunoglobulin | |
| RT | epsilon chain cDNA."; | |
| RL | Nucleic Acids Res. 11:719-726(1983). | |
| RN | [2] | |
| RP | SEQUENCE FROM N.A. | |
| RX | MEDLINE=63001945; PubMed=6288268; | |
| RA | Max E.E., Battey J., Ney R., Kirsch I.R., Leder P.; | |
| RT | "Duplication and deletion in the human immunoglobulin epsilon genes."; | |
| RN | [3] | |
| RP | SEQUENCE FROM N.A. | |
| RX | MEDLINE=84236029; PubMed=6234164; | |
| RA | Flanagan J.G., Rabbitts T.H.; | |
| RT | "The sequence of a human immunoglobulin epsilon heavy chain constant | |
| RT | region gene, and evidence for three non-allelic genes."; | |
| RL | EMBO J. 1:655-660(1982). | |
| RN | [4] | |
| RP | SEQUENCE FROM N.A. | |
| RX | MEDLINE=84207910; PubMed=6327276; | |
| RA | Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.; | |
| RT | "Long terminal repeat-like elements flank a human immunoglobulin | |
| RT | epsilon pseudogene that lacks introns."; | |
| RL | EMBO J. 1:1539-1544(1982). | |
| RN | [5] | |
| RP | PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND). | |
| RA | Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.; | |
| RT | (In) Bach M.K. (eds.). | |
| RL | Immediate hypersensitivity: modern concepts and developments, pp.1-36, | |
| RL | Marcel Dekker, New York (1978). | |
| RN | [6] | |
| RP | SEQUENCE OF 1-40; 68-114 AND 427-428 FROM N.A. | |
| RX | MEDLINE=83065234; PubMed=6815656; | |
| RA | Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J., | |
| RA | Bell L.O., Gould H.J.; | |
| RT | "Cloning and sequence determination of the gene for the human | |
| RT | immunoglobulin epsilon chain expressed in a myeloma cell line."; | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982). | |
| RN | [7] | |

| | | | | | |
|----|------|------|------|--------------|--------------------|
| 34 | 43 | 30.5 | 1581 | 1 ARO1_PNECA | Q12659 p pentafunc |
| 35 | 42.5 | 30.1 | 259 | 1 RPOD_PVRAB | Q9V198 pyrococcus |
| 36 | 42.5 | 30.1 | 332 | 1 YIE2_HSVB4 | Q02484 bovine herp |
| 37 | 42.5 | 30.1 | 550 | 1 SYM_CHLMU | Q9P107 chlamydia m |
| 38 | 42.5 | 30.1 | 611 | 1 FETA_RAT | P02773 rattus norv |
| 39 | 42.5 | 30.1 | 982 | 1 RPO1_KLULA | P05472 kluyveromyc |
| 40 | 42 | 29.8 | 248 | 1 Y4UH_RHISN | Q53200 rhizobium s |
| 41 | 42 | 29.8 | 264 | 1 YRBF_HAEIN | P34501 haemophilus |
| 42 | 42 | 29.8 | 275 | 1 TRUA_SYNY3 | P73295 synechocyst |
| 43 | 42 | 29.8 | 305 | 1 CDSA_CHLTR | O84457 c phosphati |
| 44 | 42 | 29.8 | 340 | 1 ALC2_HUMAN | P01877 homo sapien |
| 45 | 42 | 29.8 | 373 | 1 MTF2_MOUSE | Q02395 mus musculu |

3D-STRUCTURE MODELING.
 Padlan E.A., Davies D.R.;
 "A model of the F-C- of immunoglobulin E";
 Submitted (JUL-1993) to the PDB data bank.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: L00022; AAB59424.1; ALT_INIT.
 PIR: A02142; EHHU.
 PIR: A22771; A22771.
 PIR: A23195; A23195.
 PIR: I1GE; 15-JUL-92.
 MIN: 147180; .
 InterPro: IPR003006; Ig_MHC.
 InterPro: IPR003597; Ig_c1.
 Pfam: PF00047; Ig; 4; 4.
 SMART: SM00407; IGcl; 4.
 PROSITE: PS00290; IG_MHC; 3.
 Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 3D-structure.
 NON_TER 1 1
 DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
 DISULFID 15 105
 DISULFID 29 85
 DISULFID 121 121 INTERCHAIN (WITH A HEAVY CHAIN).
 DISULFID 135 193
 DISULFID 309 209 INTERCHAIN (WITH A HEAVY CHAIN).
 DISULFID 339 299
 DISULFID 345 405
 CARBOHYD 21 21 N-LINKED (GLCNAC. . .).
 CARBOHYD 49 49 N-LINKED (GLCNAC. . .).
 CARBOHYD 99 99 N-LINKED (GLCNAC. . .).
 CARBOHYD 146 146 N-LINKED (GLCNAC. . .).
 CARBOHYD 252 252 N-LINKED (GLCNAC. . .).
 CARBOHYD 275 275 N-LINKED (GLCNAC. . .).
 VARIANT 359 W -> L (IN REF. 2, POSSIBLY DUE TO
 POLYMORPHISM).
 /FTid=VAR_003885.
 SEQUENCE 428 AA; 47019 MW; 25C4CA072AA558A0 CRC64;

Query Match 60.3%; Score 85; DB 1; Length 428;
 Best Local Similarity 65.2%; Pred. No. 3.3e-05;
 Matches 15; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 2 GETYSRVTHPLPKDIVRSIAK 24
 II II IIIIIIIIIIIII
 Db 294 GETYQCRVTHPLPALMRSTK 316

 RESULT 2
 EPC_RAT STANDARD; PRT; 429 AA.
 AC P01855;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1995 (Rel. 38, Last annotation update)
 DE IG EPSILON CHAIN C REGION.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).
 RC STRAIN=LOU/C/WSL;
 RX MEDLINE=83064537; PubMed=6292865;
 RA Hellman L., Pettersson U., Engstrom A., Karlsson T., Bennich H.;

RT "Structure and evolution of the heavy chain from rat immunoglobulin
 E";
 RL Nucleic Acids Res. 10:6041-6049(1982).
 RN [2]
 RP SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).
 RX MEDLINE=83182019; PubMed=6820340;
 RA Kindsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;
 RT "A cloned cDNA probe for rat immunoglobulin epsilon heavy chain:
 construction, identification, and DNA sequence";
 RL DNA 1:335-343(1982).
 RN [3]
 RP SEQUENCE OF 205-306 FROM N.A.
 RX MEDLINE=82174576; PubMed=6803238;
 RA Hellman L., Pettersson U., Bennich H.;
 RT "Characterization and molecular cloning of the mRNA for the heavy
 (epsilon) chain of rat immunoglobulin E";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:1264-1268(1982).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: J00744; AAA41379.1; ALT_INIT.
 PIR: A02143; EHRT.
 InterPro: IPR003006; Ig_MHC.
 InterPro: IPR003597; Ig_c1.
 Pfam: PF00047; Ig; 4; 4.
 SMART: SM00407; IGcl; 1.
 SMART: SM00410; IG-like; 3.
 PROSITE: PS00290; IG_MHC; 3.
 Immunoglobulin domain; Immunoglobulin C region.
 NON_TER 1 1
 CONFLICT 168 168 R -> N (IN REF. 2).
 CONFLICT 308 308 P -> L (IN REF. 2).
 SEQUENCE 429 AA; 48671 MW; D2970B34EF8A72B0 CRC64;
 Query Match 58.2%; Score 82; DB 1; Length 429;
 Best Local Similarity 69.6%; Pred. No. 9.3e-05;
 Matches 16; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 Qy 2 GETYSRVTHPLPKDIVRSIAK 24
 II II IIIIIIIIIIIII
 Db 284 GETYQCRVDHPHPKPIVRSITK 306

 RESULT 3
 EPC_MOUSE STANDARD; PRT; 421 AA.
 AC P06336; P01856;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE IG EPSILON CHAIN C REGION.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84236092; PubMed=6329728;
 RA Ishida N., Ueda S., Hayashida H., Miyata T., Honjo T.;
 RT "The nucleotide sequence of the mouse immunoglobulin epsilon gene:
 comparison with the human epsilon gene sequence";
 RL EMBO J. 1:1117-1123(1982).
 RN [2]
 RP REVISIONS.
 RA Honjo T.;

RL Submitted (APR-1986) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE OF 34-421 FROM N.A.
 RX MEDLINE=8311774; PubMed=6818553;
 RL Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;
 RT "Cloning and nucleotide sequence of mouse immunoglobulin epsilon
 chain cDNA";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856(1982).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X01857; CAA25977.1; -
 DR EMBL; X01857; CAA25978.1; -
 DR PIR; A02145; EHMSS.
 DR PIR; A02144; EHMSS.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003597; Ig_cl.
 DR InterPro: IPR003600; Ig_like.
 DR Pfam; PF00047; ig; 4.
 DR SMART; SM00407; Igcl; 2.
 DR SMART; SM00410; Ig_like; 2.
 DR PROSITE; PS00290; IG_MHC; 3.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.

FT NON_TER 1 1
 FT DOMAIN 1 90 CH1.
 FT DOMAIN 91 197 CH2.
 FT DOMAIN 198 304 CH3.
 FT DOMAIN 305 421 CH4.
 FT DISULFID 23 75 BY SIMILARITY.
 FT DISULFID 121 180 BY SIMILARITY.
 FT DISULFID 226 285 BY SIMILARITY.
 FT DISULFID 330 392 BY SIMILARITY.
 FT CARBOHYD 43 43 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 72 72 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 84 84 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 166 166 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 365 365 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 421 AA; 47320 MW; 8F909E1F30A06B47 CRC64;

Query Match 41.1%; Score 58; DB 1; Length 421;
 Best Local Similarity 56.5%; Pred. No. 0.32;
 Matches 13; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 2 GETYYSRTHPLPKDIVRSIAK 24

DB 280 GYGQCIVDHPDFPKPIVRSITK 302

RESULT 4

YK14_YEAST
 ID YK14_YEAST STANDARD; PRT; 116 AA.
 AC P36078;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 29, Last annotation update)
 DE HYPOPHETICAL 13.6 KDA PROTEIN IN MDH1-VMA5 INTERGENIC REGION.
 GN YK1084W.

OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Pohl T.M., Pohl F.M.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; Z28084; CAA81922.1; -
 DR PIR; S37909; S37909.
 DR SGD; S0001567; YKL084W.
 KW Hypothetical protein.
 SQ SEQUENCE 116 AA; 13641 MW; 57464E5FD9A591F8 CRC64;

Query Match 38.3%; Score 54; DB 1; Length 116;
 Best Local Similarity 52.6%; Pred. No. 0.33;
 Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 7 SRVTHPLPKDIVRSIAK 25

DB 15 SRCVHHLPKDVIAIRFKC 33

RESULT 5

FETA_MOUSE
 ID FETA_MOUSE STANDARD; PRT; 605 AA.
 AC P02772;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE ALPHA-FETOPROTEIN PRECURSOR (ALPHA-FETOGLOBULIN) (ALPHA-1-
 DE FETOPROTEIN).
 DE GN AFP.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=81197641; PubMed=6164927;
 RA Law S.W., Dugaiczky A.;
 RT "Homology between the primary structure of alpha-fetoprotein, deduced
 RT from a complete cDNA sequence, and serum albumin.";
 RL Nature 291:201-205(1981).
 RN [2]
 RP REVISION TO 598.
 RX MEDLINE=88216123; PubMed=2452956;
 RA Minghetti P.P., Law S.W., Dugaiczky A.;
 RT "The rate of molecular evolution of alpha-fetoprotein approaches that
 RT of pseudogenes.";
 RL Mol. Biol. Evol. 2:347-358(1985).
 RN [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=82103990; PubMed=6172714;
 RA Eiferman F.A., Young P.R., Scott R.W., Tilghman S.M.;
 RT "Intragenic amplification and divergence in the mouse
 RT alpha-fetoprotein gene.";
 RL Nature 294:713-718(1981).
 RN [4]
 RN SEQUENCE OF 15-605 FROM N.A.
 RX MEDLINE=81117287; PubMed=6161929;
 RA Gorin M.B., Cooper D.L., Eiferman F.A., van de Rijn P., Tilghman S.M.;
 RT "The evolution of alpha-fetoprotein and albumin. I. A comparison of
 RT the primary amino acid sequences of mammalian alpha-fetoprotein and
 RT albumin.";
 RL J. Biol. Chem. 256:1954-1959(1981).
 CC -!- FUNCTION: BINDS ESTROGENS, FATTY ACIDS AND METALS.
 CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -1- TISSUE SPECIFICITY: PLASMA.
CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
CC -1- PTM: GLYCOSYLATED; HAS TWO CARBOHYDRATE CHAINS.
CC -1- PTM: SULFATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00743; CAA24118.1; -
DR EMBL; M16394; AAA37189.1; -
DR EMBL; M16394; AAA37189.1; JOINED.
DR EMBL; M16381; AAA37189.1; JOINED.
DR EMBL; M16382; AAA37189.1; JOINED.
DR EMBL; M16383; AAA37189.1; JOINED.
DR EMBL; M16384; AAA37189.1; JOINED.
DR EMBL; M16385; AAA37189.1; JOINED.
DR EMBL; M16386; AAA37189.1; JOINED.
DR EMBL; M16387; AAA37189.1; JOINED.
DR EMBL; M16388; AAA37189.1; JOINED.
DR EMBL; M16389; AAA37189.1; JOINED.
DR EMBL; M16390; AAA37189.1; JOINED.
DR EMBL; M16391; AAA37189.1; JOINED.
DR EMBL; M16392; AAA37189.1; JOINED.
DR EMBL; M16393; AAA37189.1; JOINED.
DR PIR; A03235; FPMs.
DR HGSP; P02768; IUOR.
DR MGD; MGI-87951; Afp.
DR InterPro; IPI000264; Serum_albumin.
DR Pfam; PF00273; transport_prot; 3.
DR PRINTS; PRO0002; SERUMALBUMIN.
DR PRINTS; PRO0003; AFETOPROTEIN.
DR SMART; SM00103; ALBUMIN; 3.
DR PROSITE; PS00212; ALBUMIN; 2.
DR Glycoprotein; Sulfation; Albumin; Plasma; Embryo; Repeat;
KW Metal-binding; Signal
FT SIGNAL 1 18 PROBABLE.
FT CHAIN 19 605 ALPHA-FETOPROTEIN.
FT REPEAT 28 198 1.
FT REPEAT 217 390 2.
FT REPEAT 409 588 3.
FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 498 498 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 95 110
FT DISULFID 109 120
FT DISULFID 144 189
FT DISULFID 188 197
FT DISULFID 220 266
FT DISULFID 265 273
FT DISULFID 285 299
FT DISULFID 298 309
FT DISULFID 380 389
FT DISULFID 412 458
FT DISULFID 457 468
FT DISULFID 481 497
FT DISULFID 496 507
FT DISULFID 534 579
FT DISULFID 578 587
FT CONFLICT 533 536 LCOA -> RAKL (IN REF. 4).
FT SEQUENCE 605 AA; 67337 MW; CE09E9F50D74619A CRC64;

Query Match 36.5%; Score 51.5; DB 1; Length 605;
Best Local Similarity 57.9%; Pred. No. 4.3;
Matches 11; Conservative 3; Mismatches 4; Indels 1; Gaps 1;
QY YSRVTHPLPKDIVRSTAK 24
||| |||:: |||

Db 354 YSR-THPNLPVSVILRIAK 371
RESULT 6
Y116_ADE02 STANDARD; PRT; 106 AA.
ID Y116_ADE02
AC P03287;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-APR-1990 (Rel. 14, Last annotation update)
DE HYPOTHETICAL 11.6 KDA EARLY PROTEIN.
OS Human adenovirus type 2.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10515;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83056843; PubMed=7142161;
RA Gingers T.R., Sciaky D., Gellinas R.E., Bing-Dong J., Yen C.E.,
RA Kelly M.M., Bullock P.A., Parsons B.L., O'Neill K.E., Roberts R.J.;
RT "Nucleotide sequences from the adenovirus-2 genome.";
RL J. Biol. Chem. 257:13475-13491(1982).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=83056844; PubMed=7142162;
RA Alestrom P., Akusjærvi G., Pettersson M., Pettersson U.;
RT "DNA sequence analysis of the region encoding the terminal protein
and the hypothetical N-gene product of adenovirus type 2.";
RL J. Biol. Chem. 257:13492-13498(1982).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J01917; -; NOT_ANNOTATED_CDS.
DR PIR; A03859; A03859.
DR Hypothetical protein; Early protein.
KW SEQUENCE 106 AA; 11668 MW; 12692EF62B02E3A3 CRC64;
Query Match 36.2%; Score 51; DB 1; Length 106;
Best Local Similarity 62.5%; Pred. No. 0.83;
Matches 10; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 6 YSRVTHPLPKDIVRS 21
||| |||:: |||
Db 23 YSRATHRSKGVTVRS 38
RESULT 7
GVPC_HALME STANDARD; PRT; 381 AA.
ID GVPC_HALME
AC Q02228;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE GAS VESICLE PROTEIN C.
GN GVPC.
OS Halobacterium mediterranei (Haloflex mediterranei).
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Haloflex.
OX NCBI_TaxID=2252;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DSM 1411;
RA MEDLINE=93021102; PubMed=1404376;
RA Engler C., Krueger K., Offner S., Pfeiffer F.;
RT "Three different but related gene clusters encoding gas vesicles in
RT halophilic archaea";
RL J. Mol. Biol. 227:586-592(1992).
CC -1- FUNCTION: MAY CONFER STABILITY TO THE GAS VESICLE MEMBRANES. GAS

CC VESICLES ARE SMALL, HOLLOW, GAS FILLED PROTEIN STRUCTURES THAT ARE
CC FOUND IN SEVERAL PROKARYOTIC PLANKTONIC MICROORGANISMS. THEY ALLOW
CC THE POSITIONING OF THE BACTERIA, AT THE FAVORABLE DEPTH FOR
CC GROWTH.
CC
CC -!- SUBCELLULAR LOCATION: BINDS TO THE EXTERNAL SURFACE OF THE GAS
CC VESICLE MEMBRANE.
CC
CC -!- SIMILARITY: VERY DISTANT, TO GAS VESICLE PROTEINS TYPE C OF
CC CYANOBACTERIA.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; X64701; CAA45944.1; -
CC PIR; S28115; S28115.
CC
CC Gas vesicle; Repeat.
CC
CC FT DOMAIN 22 274 7 X APPROXIMATE TANDEM REPEATS.
CC REPEAT 22 59 1.
CC FT REPEAT 60 84 2.
CC FT REPEAT 85 122 3.
CC FT REPEAT 123 160 4.
CC FT REPEAT 161 192 5.
CC FT REPEAT 193 232 6.
CC FT REPEAT 233 274 7.
CC FT DOMAIN 246 335 ASP/GLU-RICH (ACIDIC).
CC
CC SQ SEQUENCE 381 AA; 42653 MW; 9FB4819D0305921 CRC64;

Query Match 35.1%; Score 49.5; DB 1; Length 381;
Best Local Similarity 64.3%; Pred. No. 5.3;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

Qy 1 CGEYYYSRVTHPHL 14
| | | | | : | | | |
Db 343 CGE-YYQAITPHL 355

RESULT 8
GVCL_HALN1 STANDARD; PRT; 382 AA.
AC P24574; O9H117;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE GAS VESICLE PROTEIN C.1.
GN (GVPC1 OR GVPC OR VNG50326) AND (GVPC2 OR VNG6031G).
OS Halobacterium sp. (strain NRC-1), and
OS Halobacterium sp. (strain NRC-817).
OG Plasmid pNRC100, Plasmid pNRC200, and Plasmid pHH1.
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC Halobacterium.
OX NCBI_TaxID=64091, 148370;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=NRC-1; PLASMID=pNRC100;
RX MEDLINE=91323716; PubMed=1864501;
RA Jones J.G., Young D.C., Dassarma S.;
RT "Structure and organization of the gas vesicle gene cluster on the
RT Halobacterium halobium plasmid pNRC100.";
RL Gene 102:117-122(1991).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=NRC-1; PLASMID=pNRC100;
RX MEDLINE=99063795; PubMed=9847077;
RA Ng W.V., Clifo S.A., Smith T.M., Bumgarner R.E., Baskin D., Faust J.,
RA Hall B., Loretz C., Seto J., Slagel J., Hood L., Dassarma S.;
RT "Snapshot of a large dynamic replicon in a halophilic archaeon:
RT megaplasmid or minichromosome?";
RL Genome Res. 8:1131-1141(1998).

RN SEQUENCE FROM N.A.
RP STRAIN=NRC-1; PLASMID=pNRC200;
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Madden D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
[4]
RN SEQUENCE FROM N.A.
RP STRAIN=NRC-817; PLASMID=pHH1;
RX MEDLINE=92065812; PubMed=1956294;
RA Horne M., Englert C., Wimmer C., Pfeifer F.;
RT "A DNA region of 9 kbp contains all genes necessary for gas vesicle
RT synthesis in halophilic archaeobacteria.";
RL Mol. Microbiol. 5:1159-1174(1991).
[5]
RN SEQUENCE FROM N.A.
RP STRAIN=NRC-817; PLASMID=pHH1;
RX MEDLINE=93021102; PubMed=1404376;
RA Englert C., Krueger K., Offner S., Pfeifer F.;
RT "Three different but related gene clusters encoding gas vesicles in
RT halophilic archaea.";
RL J. Mol. Biol. 227:586-592(1992).
CC -!- FUNCTION: MAY CONFER STABILITY TO THE GAS VESICLE MEMBRANES. GAS
CC VESICLES ARE SMALL, HOLLOW, GAS FILLED PROTEIN STRUCTURES THAT ARE
CC FOUND IN SEVERAL PROKARYOTIC PLANKTONIC MICROORGANISMS. THEY ALLOW
CC THE POSITIONING OF THE BACTERIA, AT THE FAVORABLE DEPTH FOR
CC GROWTH.
CC
CC -!- SUBCELLULAR LOCATION: BINDS TO THE EXTERNAL SURFACE OF THE GAS
CC VESICLE MEMBRANE.
CC
CC -!- SIMILARITY: VERY DISTANT, TO GAS VESICLE PROTEINS TYPE C OF
CC CYANOBACTERIA.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; M58557; AAA98197.1; -
CC EMBL; AF016485; AAC82810.1; -
CC EMBL; AE005142; AAG20727.1; -
CC EMBL; X57161; CAA40451.1; -
CC EMBL; X64729; CAA45981.1; -
CC PIR; J01122; J01122.
CC PIR; S15483; S15483.
CC PIR; S28128; S28128.
CC
CC Gas vesicle; Plasmid; Repeat; Complete proteome.
CC
CC FT DOMAIN 22 284 7 X APPROXIMATE TANDEM REPEATS.
CC REPEAT 22 60 1.
CC FT REPEAT 61 92 2.
CC FT REPEAT 93 130 3.
CC FT REPEAT 131 168 4.
CC FT REPEAT 169 200 5.
CC FT REPEAT 201 240 6.
CC FT REPEAT 241 284 7.
CC FT DOMAIN 254 336 ASP/GLU-RICH (ACIDIC).
CC
CC SQ SEQUENCE 382 AA; 42391 MW; 171DBEB4C0364F46 CRC64;

Query Match 35.1%; Score 49.5; DB 1; Length 382;
Best Local Similarity 64.3%; Pred. No. 5.3;
Matches 9; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

```
Qy 1 CGETYYSRVTHPL 14
    ||| || : |||
Db 344 CGE-YYQAITERHL 356
    STANDARD; PRT; 384 AA.

RESULT 9
ID GBA2_PEA
AC O04279;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE GUANINE NUCLEOTIDE-BINDING PROTEIN ALPHA-2 SUBUNIT (GP-ALPHA-2).
GN GBA2 OR GA2.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ALASKA; TISSUE=Apical bud;
RA Marsh J.F., Kaufman L.S.;
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 2 (G(S)).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U97044; AAB57826.1; -
DR InterPro: IPR001019; Gprotein_alpha.
DR Pfam: PF00503; G-alpha.1.
DR PRINTS: PR00318; GPROTEIN.
DR SMART: SM00275; G-alpha.1.
DR GTP-binding: Transducer; ADP-ribosylation.
DR NP_BIND 46 53 GTP (BY SIMILARITY).
DR NP_BIND 219 223 GTP (BY SIMILARITY).
DR NP_BIND 288 291 GTP (BY SIMILARITY).
DR MOD_RES 191 191 ADP-RIBOSYL[1] (BY ACTION OF CTX).
SQ SEQUENCE 384 AA; 44569 MW; 6B0963393FD20BA4 CRC64;

Query Match 34.0%; Score 48; DB 1; Length 384;
Best Local Similarity 50.0%; Pred. No. 8.9;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 7 SRVTHPLPKDIVRSI 22
    ||| : ||| || : |
Db 125 SRLDYPHLLTKDLAREI 140
    STANDARD; PRT; 1161 AA.

RESULT 10
POL_SFV1
ID POL_SFV1
AC P23074;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE POL POLYPROTEIN [CONTAINS: PROTEASE (EC 3.4.23.-); REVERSE
DE TRANSCRIPTASE (EC 2.7.7.49); ENDONUCLEASE].
GN POL.
OS Simian foamy virus (type 1) (SFV-1).
```

```
OC Viruses; Retrovird viruses; Retroviridae; Spumavirus.
OX NCBI_TaxID=11964;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91276270; PubMed=1647358;
RA Kupiec J.-J., Kay A., Hayat M., Ravier R., Perles J., Galibert F.;
RT "Sequence analysis of the simian foamy virus type 1 genome.";
RL Gene 101:185-194(1991).
RN [2]
RP SEQUENCE OF 1-970 FROM N.A.
RX MEDLINE=91361538; PubMed=1653483;
RA Mergia A., Luciw P.A.;
RT "Replication and regulation of primate foamy viruses.";
RL Virology 184:475-482(1991).
RN [3]
RP SEQUENCE OF 969-1161 FROM N.A.
RX MEDLINE=90081048; PubMed=2152825;
RA Mergia A., Shaw K.E.S., Lackner J.E., Luciw P.A.;
RT "Relationship of the env genes and the endonuclease domain of the pol
RT genes of simian foamy virus type 1 and human foamy virus.";
RL J. Virol. 64:406-410(1990).
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A9.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X54482; -; NOT ANNOTATED_CDS.
DR EMBL: X58484; CA41394.1; -
DR EMBL: M33561; AAA47793.1; -
DR PIR: A33562; A33562.
DR PIR: S15566; S15566.
DR PIR: S18738; S18738.
DR HSSP: P03355; IMWL.
DR MEROPS: A09.001; -
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR000477; RVTse.
DR InterPro: IPR001584; Rve.
DR InterPro: IPR001641; Spuma_A9ptase.
DR Pfam: PF00075; InaseH; 1.
DR Pfam: PF00665; rve; 1.
DR Pfam: PF00078; rvt; 1.
DR PRINTS: PR00920; SPUMVIRPTASE.
DR Transferase; RNA-directed DNA polymerase; Hydrolase; Nuclease;
DR Endonuclease; Aspartyl protease; Polyprotein.
FT ACT_SITE 36 36 BY SIMILARITY.
FT CONFLICT 236 236 T -> I (IN REF. 2).
FT CONFLICT 910 910 S -> G (IN REF. 2).
FT CONFLICT 950 950 A -> T (IN REF. 2).
SQ SEQUENCE 1161 AA; 131717 MW; 08D09D1D6BB4C14C CRC64;

Query Match 34.0%; Score 48; DB 1; Length 1161;
Best Local Similarity 47.4%; Pred. No. 28;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Qy 7 SRVTHPLPKDIVRSIAC 25
    | : | | : | | : |
Db 840 SKYWWPNLRKDVVKSIROC 858

RESULT 11
GBAL_LYCES
ID GBAL_LYCES STANDARD; PRT; 384 AA.
AC P26981;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
```

DE GUANINE NUCLEOTIDE-BINDING PROTEIN ALPHA-1 SUBUNIT (GP-ALPHA-1).
GN GPAL OR GAI.
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4081;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CV, VF36; TISSUE=Pistil;
RX MEDLINE=92084110; PubMed=1748292;
RA Ma H., Yanofsky M.F., Huang H.;
RT "Isolation and sequence analysis of TGA1 cDNAs encoding a tomato G
protein alpha subunit.";
RL Gene 107:189-195(1991).
CC -1- FUNCTION: GUANINE NUCLEOTIDE-BINDING PROTEINS (G PROTEINS) ARE
CC INVOLVED AS MODULATORS OR TRANSDUCERS IN VARIOUS TRANSMEMBRANE
CC SIGNALING SYSTEMS.
CC -1- SUBUNIT: G PROTEINS ARE COMPOSED OF 3 UNITS (ALPHA, BETA & GAMMA).
CC THE ALPHA CHAIN CONTAINS THE GUANINE NUCLEOTIDE BINDING SITE.
CC -1- SIMILARITY: BELONGS TO THE G-ALPHA FAMILY. SUBFAMILY 2 (G(S)).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M74419; AAA34167.1; -
DR PIR; JH0514; RGT00A.
DR HSSP; P10824; IAS3.
DR InterPro; IPR001019; Gprotein_alpha.
DR Pfam; PF00503; G-alpha: 1.
DR PRINTS; PR00318; GPROTEINA.
DR PRINTS; PR01242; GPROTEINAPIT.
DR SMART; SM00275; G-alpha; 1.
KW GTP-binding; Transducer; ADP-ribosylation.
FT NP_BIND 46 53 GTP (BY SIMILARITY).
FT NP_BIND 219 223 GTP (BY SIMILARITY).
FT NP_BIND 288 291 GTP (BY SIMILARITY).
FT MOD_RES 191 191 ADP-RIBOSYL[1] (BY ACTION OF CTX).
SQ SEQUENCE 384 AA; 44903 MW; 5C95ED74EFC01A0A CRC64;

Query Match 33.3%; Score 47; DB 1; Length 384;
Best Local Similarity 53.3%; Pred. No. 12;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 8 RVTHPLPKDIVRSI 22
| : ||| ||: |
Db 126 RLDPHPLTKDLVQDI 140

RESULT 12
ID BAR1_HUMAN STANDARD; PRT; 777 AA.
AC Q99728; O43574;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE BRCA1-ASSOCIATED RING DOMAIN PROTEIN 1 (BARD-1).
GN BAR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=B-cell.
RX MEDLINE=97099454; PubMed=8944023;
RA Wu L.C., Wang Z.W., Tsan J.T., Spillman M.A., Phung A., Xu X.L.,

RA Yang M.-C.W., Hwang L.-Y., Bowcock A.M., Baer R.;
RT "Identification of a RING protein that can interact in vivo with the
RT BRCA1 gene product.";
RL Nat. Genet. 13:430-440(1996).
RN [2]
RN SEQUENCE FROM N.A., AND VARIANTS.
RX MEDLINE=98087570; PubMed=9425226;
RA Thai T.H., Du F., Tsan J.T., Jin Y., Phung A., Spillman M.A.,
RA Massa H.F., Muller C.Y., Ashfaq R., Mathis J.M., Miller D.S.,
RA Trask B.J., Baer R., Bowcock A.M.;
RT "Mutations in the BRCA1-associated RING domain (BARD1) gene in primary
RT breast, ovarian and uterine cancers.";
RL Hum. Mol. Genet. 7:195-202(1998).
RN [3]
RN DOMAINS.
RX MEDLINE=99150352; PubMed=10026184;
RA Meza J.E., Brzovic P.S., King M.-C., Kleit R.E.;
RT "Mapping the functional domains of BRCA1. Interaction of the ring
RT finger domains of BRCA1 and BARD1.";
RL J. Biol. Chem. 274:5659-5665(1999).
RN [4]
RN POSSIBLE FUNCTION.
RX MEDLINE=99407028; PubMed=10477523;
RA Kiehlman F.E., Manley J.L.;
RT "Functional interaction of BRCA1-associated BARD1 with polyadenylation
RT factor CstF-50.";
RL Science 285:1576-1579(1999).
CC -1- FUNCTION: IMPLICATED IN BRCA1-MEDIATED TUMOR SUPPRESSION. MAY, AS
CC PART OF THE RNA POLYMERASE-2 HOLOENZYME, FUNCTION IN THE CELLULAR
CC RESPONSE TO DNA DAMAGE. IN VITRO, INHIBITS PRE-MRNA 3' CLEAVAGE.
CC -1- SUBUNIT: HOMO-AND HETERODIMER. PREFERENTIALLY FORMS HETERODIMERS
CC WITH BRCA1. CAN ALSO FORM HETERODIMERS WITH CstF-50.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. DURING S PHASE OF THE CELL CYCLE,
CC CO-LOCALIZES WITH BRCA1 INTO DISCRETE SUBNUCLEAR FOCI.
CC -1- PTM: PROCESSED DURING APOPTOSIS. THE HOMODIMER IS MORE SUSCEPTIBLE
CC TO PROTEOLYTIC CLEAVAGE THAN THE BARD1/BRCA1 HETERODIMER.
CC -1- DISEASE: DEFECTS IN BARD1 GENE ARE FOUND IN PRIMARY BREAST,
CC OVARIAN AND UTERINE CANCERS.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
CC -1- SIMILARITY: CONTAINS 3 ANK REPEATS.
CC -1- SIMILARITY: CONTAINS 2 BRCT DOMAINS.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-26 IS THE INITIATOR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U76638; AAB38316.1; -
DR EMBL; AF038042; AAB99978.1; -
DR EMBL; AF038034; AAB99978.1; JOINED.
DR EMBL; AF038035; AAB99978.1; JOINED.
DR EMBL; AF038036; AAB99978.1; JOINED.
DR EMBL; AF038037; AAB99978.1; JOINED.
DR EMBL; AF038038; AAB99978.1; JOINED.
DR EMBL; AF038039; AAB99978.1; JOINED.
DR EMBL; AF038040; AAB99978.1; JOINED.
DR EMBL; AF038041; AAB99978.1; JOINED.
DR HSSP; P15919; IRMD.
DR MIM; 601593; -
DR InterPro; IPR002110; ANK.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00023; ank; 3.
DR Pfam; PF00533; BRCT; 2.
DR SMART; SM00248; ANK; 3.
DR SMART; SM00292; BRCT; 2.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00088; ANK_REPEAT; 3.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR

DR PROSITE; PS00172; BRCT; 2.
DR zinc-finger; ZINC_FINGER_C3HC4; 1.
KW ANK repeat.
FT 2N-FING 50 86 RING-TYPE.
FT REPEAT 427 459 ANK 1.
FT REPEAT 460 492 ANK 2.
FT REPEAT 463 525 ANK 3.
FT DOMAIN 570 653 BRCT 1.
FT DOMAIN 667 777 BRCT 2.
FT DOMAIN 26 119 REQUIRED FOR INTERACTION WITH BRCA1.
FT DOMAIN 400 403 POLY-SER.
FT DOMAIN 542 545 POLY-LEU.
FT VARIANT 24 24 P -> S (FREQUENT POLYMORPHISM IN CAUCASIANS; LESS FREQUENT IN AFRICANS).
FT VARIANT 153 153 /FTid=VAR_010354.
FT VARIANT 507 507 /FTid=VAR_010355.
FT VARIANT 557 557 /FTid=VAR_010356.
FT VARIANT 564 564 /FTid=VAR_010357.
FT VARIANT 658 658 /FTid=VAR_010358.
FT VARIANT 695 695 /FTid=VAR_010359.
FT VARIANT 761 761 /FTid=VAR_010360.
FT CONFLICT 378 378 R -> S (IN REF. 2).
FT CONFLICT 406 406 O -> R (IN REF. 2).
SQ SEQUENCE 777 AA; 86619 MW; AAEC604524E8D2B CRC64;

Query Match 33.3%; Score 47; DB 1; Length 777;
Best Local Similarity 52.6%; Pred. No. 26;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 7 SRVTHPHLPKDIVRSIAKC 25
| | | | : | | | : | |
Db 602 STVTHVVPDVGQSLKLC 620

RESULT 13
HB2D_RAT
ID HB2D_RAT STANDARD; PRT; 264 AA.
AC P18211;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1995 (Rel. 38, Last annotation update)
DE RT1 CLASS II HISTOCOMPATIBILITY ANTIGEN, D-1 BETA CHAIN PRECURSOR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LEWIS FAMILIARIS; TISSUE=Bone marrow;
RX MEDLINE=90356406; PubMed=2388838;
RA Syha-Jedelhauser J., Reske K.;
RT "Sequence of rat cDNA clone pLR beta 112 coding for the RT1.D beta 1 chain."
RL Nucleic Acids Res. 18:4598-4598(1990).
CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE IMMUNE SYSTEM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X53054; CAA37221.1; -
DR PIR; S10989; S10989.
DR HSSP; P13760; ZSEB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003593; MHC_II_beta.
DR Pfam; PF00047; Ig_1.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
DR SMART; SM00407; IG1; 1.
DR PROSITE; PS00290; IG_MHC; 1.
KW MHC II; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 26
FT CHAIN 27 264 RT1 CLASS II HISTOCOMPATIBILITY ANTIGEN, D-1 BETA CHAIN.
FT DOMAIN 27 120 EXTRACELLULAR BETA-1.
FT DOMAIN 121 215 EXTRACELLULAR BETA-2.
FT DOMAIN 216 226 CONNECTING PEPTIDE.
FT TRANSMEM 227 248
FT DOMAIN 249 264 CYTOPLASMIC TAIL.
FT DISULFID 42 106 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 264 AA; 29876 MW; 158F357355177DA1 CRC64;

Query Match 32.6%; Score 46; DB 1; Length 264;
Best Local Similarity 47.1%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 2 GEYYSRVTHPHLPKDI 18
| | | : | | | :
Db 195 GEYTCQVEHPSPV 211

RESULT 14
UVRC_PSEFL
ID UVRC_PSEFL STANDARD; PRT; 607 AA.
AC P32966;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE EXCINUCLEASE ABC SUBUNIT C.
GN UVRC.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BL915;
RX MEDLINE=94355677; PubMed=8075420;
RA Gaffney T.D., Lam S.T., Ligon J., Gates K., Frazelle A., Maio J., Hill S., Goodwin S., Torkewitz N., Allshouse A.M., Kempf H.J., Becker J.O.;
RA "Global regulation of expression of antifungal factors by a Pseudomonas fluorescens biological control strain."
RL Mol. Plant Microbe Interact. 7:455-463(1994).
RN [2]
RP SEQUENCE OF 1-232 FROM N.A.
RC STRAIN=CHAO;
RX MEDLINE=92179223; PubMed=1311842;
RA Laville J., Voisard C.P., Keel C., Maurhofer M., Difago G., Haas D.;
RA "Global control in Pseudomonas fluorescens mediating antibiotic synthesis and suppression of black root rot of tobacco."
RL Proc. Natl. Acad. Sci. U.S.A. 89:1562-1566(1992).
CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). ATTACHES TO THE

CC UVRA-UVRB COMPLEX, DISPLACING UVRA, AND THE DAMAGED DNA STRAND IS
CC NICKED ON BOTH SIDES OF THE DAMAGED SITE (BY SIMILARITY).
CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS; UVRA, UVRB AND UVRC.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE UVRC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L29642; AAA98758.1; -
CC EMBL; M80913; AAA25822.1; -
CC InterPro: IPR000445; HHH.
CC InterPro: IPR003583; HHH_1.
CC InterPro: IPR001943; UVR.
CC InterPro: IPR001162; UVR_2.
CC InterPro: IPR000305; UVR_1.
CC Pfam; PF01541; Excl_endo_N; 1.
CC Pfam; PF00633; HHH; 2.
CC Pfam; PF02151; UVR; 1.
CC ProDom; PD005870; UVR_2; 1.
CC SMART; SM00465; GIYC; 1.
CC SMART; SM00278; HHH1; 2.
CC SOS response; Excision nuclease; DNA repair.
CC CONFLICT 3 4 DP -> EQ (IN REF. 2).
CC CONFLICT 28 28 T -> A (IN REF. 2).
CC CONFLICT 41 41 S -> N (IN REF. 2).
CC CONFLICT 44 44 A -> S (IN REF. 2).
CC CONFLICT 52 52 L -> Q (IN REF. 2).
CC CONFLICT 61 61 G -> A (IN REF. 2).
CC CONFLICT 111 111 E -> D (IN REF. 2).
CC CONFLICT 123 123 K -> R (IN REF. 2).
CC CONFLICT 184 184 A -> P (IN REF. 2).
CC CONFLICT 211 212 AG -> TA (IN REF. 2).
CC CONFLICT 215 215 Q -> A (IN REF. 2).
CC SEQUENCE 607 AA; 67302 MW; 179D412EC0B4398C CRC64;

Query Match 32.6%; Score 46; DB 1; Length 607;
Best Local Similarity 40.0%; Pred. No. 28;
Matches 10; Conservative 2; Mismatches 13; Indels 0; Gaps 0;
QY 1 CGETYYSRVTHPLPKDIVRSIAK 25
DB 154 CEDSYKNTTRPCLOYQIKRCKAPC 178

RESULT 15
KORB_METJA STANDARD; PRT; 270 AA.
AC Q57957;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 2-OXOGLUTARATE SYNTHASE SUBUNIT KORB (EC 1.2.7.3) (2-KETOGLUTARATE
DE OXIDOREDUCTASE BETA CHAIN) (KOR) (2-OXOGLUTARATE-FERREDOXIN
DE OXIDOREDUCTASE BETA SUBUNIT).
GN KORB OR MJ0537.
OS Methanococcus jannaschli.
OS Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RC MEDLINE=96337999; PubMed=868087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,

RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT *Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschli.;
RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: 2-OXOBUTANOATE + COA + OXIDIZED FERREDOXIN -
CC PROPHANOYL-COA + CO(2) + REDUCED FERREDOXIN.
CC -1- SUBUNIT: HETEROTETRAMER OF THE KORA, KORB, KORC AND KORD
CC SUBUNITS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67503; AAB98531.1; -
CC TIGR; MJ0537; -
KW Oxidoreductase; Complete proteome.
SQ SEQUENCE 270 AA; 29894 MW; 4f73331CF37216554 CRC64;

Query Match 32.3%; Score 45.5; DB 1; Length 270;
Best Local Similarity 52.2%; Pred. No. 14;
Matches 12; Conservative 2; Mismatches 8; Indels 1; Gaps 1;
QY 2 GETYYYSRVTHPLPKDIVRSIAK 24
DB 164 GATYVARWTTAH-DIQLVRSIKK 185

Search completed: March 4, 2002, 13:08:42
Job time: 639 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:08:10 ; Search time 79.15 seconds
(without alignments)
46.201 Million cell updates/sec

Title: US-09-701-623C-6

Perfect score: 141

Sequence: 1 CGETYYSRVTHPLPKDIVRSIAKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------------|
| 1 | 54 | 38.3 | 394 | 5 | 017620 caenorhabdi |
| 2 | 54 | 38.3 | 1625 | 10 | 008367 zea mays (m |
| 3 | 54 | 38.3 | 2311 | 10 | 048959 triticum ae |
| 4 | 54 | 38.3 | 2325 | 10 | 041743 zea mays (m |
| 5 | 52 | 36.9 | 1756 | 10 | 080975 O80975 arabidopsis |
| 6 | 51 | 36.2 | 195 | 1 | 09VIA1 Q9VIA1 pyrococcus |
| 7 | 51 | 36.2 | 416 | 2 | 09X4C5 Q9X4C5 escherichia |
| 8 | 50 | 35.5 | 539 | 10 | 09LZX0 Q9LZX0 arabidopsis |
| 9 | 49 | 34.8 | 684 | 13 | 090544 Q90544 ginglymosto |
| 10 | 49 | 34.8 | 1095 | 4 | 075044 Q75044 homo sapien |
| 11 | 49 | 34.8 | 1865 | 10 | 043248 Q43248 zea mays (m |
| 12 | 48 | 34.0 | 141 | 12 | 087052 Q87052 simian foam |
| 13 | 48 | 34.0 | 388 | 2 | 09KY27 Q9KY27 streptomyce |
| 14 | 48 | 34.0 | 460 | 5 | 09V7W8 Q9V7W8 drosophila |
| 15 | 47.5 | 33.7 | 421 | 5 | 09BIC9 Q9BIC9 trichinella |
| 16 | 47 | 33.3 | 320 | 10 | 09IQ87 Q9IQ87 arabidopsis |
| 17 | 47 | 33.3 | 372 | 10 | 040404 Q40404 nicotiana p |
| 18 | 47 | 33.3 | 384 | 10 | 09FV62 Q9FV62 nicotiana t |
| 19 | 47 | 33.3 | 518 | 13 | 09W6R2 Q9W6R2 fugu rubrip |

| | | | | | |
|----|------|------|------|----|--------|
| 20 | 47 | 33.3 | 542 | 3 | 09HFC4 |
| 21 | 47 | 33.3 | 822 | 10 | 09FNM8 |
| 22 | 47 | 33.3 | 2254 | 10 | 09LNO2 |
| 23 | 47 | 33.3 | 2257 | 10 | 040326 |
| 24 | 46.5 | 33.0 | 248 | 10 | 09L2D6 |
| 25 | 46.5 | 33.0 | 550 | 5 | 017145 |
| 26 | 46 | 32.6 | 198 | 1 | 059359 |
| 27 | 46 | 32.6 | 245 | 7 | 031270 |
| 28 | 46 | 32.6 | 261 | 7 | 09TQA7 |
| 29 | 46 | 32.6 | 264 | 7 | 09TQA5 |
| 30 | 46 | 32.6 | 382 | 2 | 09S2V1 |
| 31 | 46 | 32.6 | 492 | 10 | 009484 |
| 32 | 46 | 32.6 | 1010 | 10 | 09ZPN1 |
| 33 | 46 | 32.6 | 1205 | 10 | 09ZNV6 |
| 34 | 46 | 32.6 | 1251 | 10 | 09SSY0 |
| 35 | 45.5 | 32.3 | 213 | 2 | 09FC29 |
| 36 | 45.5 | 32.3 | 342 | 5 | 09U1B1 |
| 37 | 45.5 | 32.3 | 407 | 5 | 09VTB0 |
| 38 | 45.5 | 32.3 | 474 | 2 | 069214 |
| 39 | 45.5 | 32.3 | 603 | 10 | 09LVE3 |
| 40 | 45 | 31.9 | 206 | 2 | 044322 |
| 41 | 45 | 31.9 | 242 | 12 | 039842 |
| 42 | 45 | 31.9 | 242 | 12 | 089465 |
| 43 | 45 | 31.9 | 242 | 12 | 067390 |
| 44 | 45 | 31.9 | 242 | 12 | 09WAE1 |
| 45 | 45 | 31.9 | 242 | 12 | 09WAD3 |

ALIGNMENTS

RESULT 1

017620 PRELIMINARY; PRT; 394 AA.
AC 017620;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)
DE C29F7.1 PROTEIN.
GN C29F7.1
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RA Mcmurray A.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shawneen R.,
RA Smalton N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans *;
RL Nature 368:32-38(1994).
DR EMBL; Z92827; CAB07326.1;
SQ SEQUENCE 394 AA; 44297 MW; D4424A972AE24451 CRC64;

Query Match 38.3%; Score 54; DB 5; Length 394;
Best Local Similarity 60.0%; Pred. No. 6.9;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 11 HPHLPKDIVRSIAKC 25

```
Db 62 HPNLPKNVWIKIASC 76
||:||||:| || |
RESULT 2
ID Q08367 PRELIMINARY; PRT; 1625 AA.
AC Q08367;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 17, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE ACETYL-COA CARBOXYLASE (EC 6.4.1.2) (ACC)
[CONTAINS: BIOTIN CARBOXYLASE (EC 6.3.4.14)] (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
SQ SEQUENCE FROM N.A.
STRAIN-B73; TISSUE=LEAF;
RX MEDLINE=94154342; PubMed=7906562;
RA Ashton A.R., Jenkins C.L., Whitfield P.R.;
RT "Molecular cloning of two different cDNAs for maize acetyl CoA
carboxylase."
RL Plant Mol. Biol. 24:35-49(1994).
CC -!- FUNCTION: THIS PROTEIN CARRIES THREE CATALYTIC FUNCTIONS: BIOTIN
CARBOXYL CARRIER PROTEIN, BIOTIN CARBOXYLASE, AND
CARBOXYLTRANSFERASE.
CC -!- CATALYTIC ACTIVITY: ATP + ACETYL-COA + HCO(3)(-) = ADP +
ORTHOPOSPHATE + MALONYL-COA.
CC -!- COFACTOR: BIOTIN.
CC -!- ENZYME REGULATION: BY PHOSPHORYLATION.
CC -!- PATHWAY: FIRST STEP (RATE LIMITING) IN LONG-CHAIN FATTY ACID
SYNTHESIS.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: PARTIAL TO CARBAMYL PHOSPHATE SYNTHETASES.
DR EMBL; 224449; CAA80822.1;
DR Mendel; 16199; Zeama:1004;16199.
DR InterPro; IPR000089; Biotin_lipoyl.
DR Pfam; PF00364; Biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Fatty acid biosynthesis; Biotin; Ligase; Multifunctional enzyme.
FT NON_TER 1 1
SQ SEQUENCE 1625 AA; 181392 MW; 3624EB042D1FF6B0 CRC64;

Query Match 38.3%; Score 54; DB 10; Length 1625;
Best Local Similarity 55.6%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYSRVTHPHLPKDVR 20
||| ||: ||| ||:
Db 459 ETYISLYPHLVKDSIQ 476

RESULT 3
ID O48959 PRELIMINARY; PRT; 2311 AA.
AC O48959;
DT 01-JUN-1993 (TREMREL. 06, Created)
DT 01-JUN-1993 (TREMREL. 06, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE ACETYL-COENZYME A CARBOXYLASE.
OS ACC-1.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]

SEQUENCE FROM N.A.
STRAIN=CV, HARD RED WINTER TAM 107;
RX MEDLINE=98054381; PubMed=9391173;
RA Gornicki P., Faris J., Podkowinski J., Gill B., Haselkorn R.;
RT "Plastid-localized acetyl-CoA carboxylase of bread wheat is encoded by
a single gene on each of the three ancestral chromosome sets."
RL Proc. Natl. Acad. Sci. U.S.A. 94:14179-14184(1997).
DR EMBL; AF029895; AAC39330.1;
DR EMBL; AF029896; AAC39331.1;
DR HSSP; P24182; 1DV1.
DR Mendel; 28005; Triae:1004;28005.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR InterPro; IPR000901; CPSase.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
SQ SEQUENCE 2311 AA; 254970 MW; 1BBC5CDD0F9F984B CRC64;

Query Match 38.3%; Score 54; DB 10; Length 2311;
Best Local Similarity 55.6%; Pred. No. 43;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYSRVTHPHLPKDVR 20
||| ||: ||| ||:
Db 1159 ETYISRLYQPHLVKDSIQ 1176

RESULT 4
Q41743
ID Q41743 PRELIMINARY; PRT; 2325 AA.
AC Q41743;
DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE ACETYL-COENZYME A CARBOXYLASE.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
SQ SEQUENCE FROM N.A.
STRAIN=AL88; TISSUE=LEAF;
RX MEDLINE=95557420; PubMed=7630949;
RA Egli M.A., Lutz S.M., Somers D.A., Gengenbach B.G.;
RT "A maize acetyl-coenzyme A carboxylase cDNA sequence."
RL Plant Physiol. 108:1299-1300(1995).
DR EMBL; U19183; AAA80214.1;
DR HSSP; P24182; 1DV1.
DR Mendel; 11125; Zeama:1004;11125.
DR InterPro; IPR000089; Biotin_lipoyl.
DR InterPro; IPR000022; Carboxyl_trans.
DR InterPro; IPR000901; CPSase.
DR Pfam; PF00364; biotin_lipoyl; 1.
DR Pfam; PF01039; Carboxyl_trans; 1.
DR Pfam; PF00289; CPSase_L_chain; 2.
DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
SQ SEQUENCE 2325 AA; 257151 MW; ECB0AEF3FE26FE0C CRC64;

Query Match 38.3%; Score 54; DB 10; Length 2325;
Best Local Similarity 55.6%; Pred. No. 43;
Matches 10; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 ETYSRVTHPHLPKDVR 20
||| ||: ||| ||:
Db 1158 ETYISRLYQPHLVKDSIQ 1175

RESULT 5
```

080975 ID O80975 PRELIMINARY; PRT; 1756 AA.
AC O80975;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE F26C24.9 PROTEIN.
GN F26C24.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
RA Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,
RA Somerville C.R., Venter J.C.;
RT "Arabidopsis thaliana chromosome II BAC F26C24 genomic sequence.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AAC04705; AAC24188.1; -
DR InterPro: IPR003653; SUMO_protease.
SQ SEQUENCE 1756 AA; 197367 MW; 68FC3B5FCC282775 CRC64;

Query Match 36.9%; Score 52; DB 10; Length 1756;
Best Local Similarity 37.5%; Pred. No. 63;
Matches 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPHLPKDIVRSIAK 24
DB 1337 CRNTLLNMTNPINPDLLAKIAK 1360
| | :|:|:| | :|:|:| |
| | :|:|:| | :|:|:| |

RESULT 6
Q9VIA1 ID Q9VIA1 PRELIMINARY; PRT; 195 AA.
AC Q9VIA1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 21.8 KDA PROTEIN.
GN PAB0359.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure and evolution";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ248284; CAB49448.1; -
DR InterPro: IPR000051; SAM_bind.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 195 AA; 21778 MW; 471178A62097C02C CRC64;

Query Match 36.2%; Score 51; DB 1; Length 195;
Best Local Similarity 47.8%; Pred. No. 9.1;
Matches 11; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

QY 2 GETYSRVTHP--HLPKDIVRSI 22
DB 119 GERFHSITNPVPHAGKDIUREI 141
| | :|:|:| | :|:|:| |
| | :|:|:| | :|:|:| |

RESULT 7
Q9X4C5 ID Q9X4C5 PRELIMINARY; PRT; 416 AA.
AC Q9X4C5;

DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE K30 CAPSULE BIOSYNTHESIS CLUSTER, PARTIAL SEQUENCE.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E69;
RX MEDLINE=99217011; PubMed=10200954;
RA Drummelsmith J., Whitfield C.;
RT "Gene products required for surface expression of the capsular form of the group 1 K antigen in Escherichia coli (O9a:K30).";
RL Mol. Microbiol. 31:1321-1332(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=E69;
RA Drummelsmith J., Whitfield C.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF104912; AAD21570.2; -
SQ SEQUENCE 416 AA; 47432 MW; 1C903B56E52EFF20 CRC64;

Query Match 36.2%; Score 51; DB 2; Length 416;
Best Local Similarity 52.4%; Pred. No. 20;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 4 TYYSRVTHPHLPKDIVRSIAK 24
DB 384 TYEKGVTHTFVVDKIRISIK 404
| | :|:|:| | :|:|:| |
| | :|:|:| | :|:|:| |

RESULT 8
Q9LZX0 ID Q9LZX0 PRELIMINARY; PRT; 539 AA.
AC Q9LZX0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 59.4 KDA PROTEIN.
GN T20LJ5_30.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Terry N., Ardiles W., Buysschaert C., Dasseville R.,
RA De Clerck R., De Keyser A., Neyt P., Rouze P., Van Den Daele H.,
RA Villarroel R., Gielen J., Van Montagu M., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162351; CAB82746.1; -
DR InterPro: IPR002014; HRS.
DR Pfam; PF00790; VHS; 1.
DR ProDom; PD003686; HRS; 1.
DR SMART; SM00288; VHS; 1.
KW Hypothetical protein.
SQ SEQUENCE 539 AA; 59412 MW; FCA23BA0BC9ECE2E CRC64;

Query Match 35.5%; Score 50; DB 10; Length 539;
Best Local Similarity 33.3%; Pred. No. 37;
Matches 8; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPHLPKDIVRSIAK 24

SQ SEQUENCE 388 AA; 42992 MW; 3AC37DADDFC59C28 CRC64;

 Query Match 34.0%; Score 48; DB 2; Length 388;
 Best Local Similarity 43.5%; Pred. No. 51;
 Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps

 QY 2 GETYSRVTHPLPKDIVRSIAK 24
 I I I I I I I I I I :
 Db 37 GATLYSPATRPSLADDILKQSR 59

 RESULT 14
 Q9VW78 PRELIMINARY; PRT; 460 AA.
 ID Q9VW78
 AC Q9VW78;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DC 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CG6796 PROTEIN.
 GN CG6796.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephyroridae; drosophilidae; Drosophila.
 ON NCBI_Taxid=7227;
 RX [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=BERKELEY.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz N.S., Gelbart W.M., Glasser K.,
 RA Folsler C., Gabrielian A.E., Gong N.S., Guan P., Harris M.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Hernandez J.R., Houck J.,
 RA Harris N.L., Harvey D., Helman T.J., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Hostin D., Houston K.A., Howland T.J., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Jalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AF003804; RAF5792.1; -;
 DR FlyBase; Fgn0034177; CG6796.
 DR InterPro; IPR002106; AA_trna_ligase_II.
 DR InterPro; IPR002309; tRNA-synt_2.
 DR InterPro; IPR002312; trna-synt-asp.

| | | | | | |
|---------------------------------------------------------------------------------------|------------------------------------------------------------------------|-------------------------|-----------|----------------|--------|
| SQ | SEQUENCE | 388 AA; | 42992 MW; | 3AC37DADEF5C28 | CRC64; |
| <hr/> | | | | | |
| Query Match 34.0%; Score 48; DB 2; Length 388; | | | | | |
| Best Local Similarity 43.5%; Pred. No. 51; | | | | | |
| Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps | | | | | |
| <hr/> | | | | | |
| QY | 2 | GETYSRVTHPHLPKDIVRSIAK | 24 | | |
| | | : | | | |
| DB | 37 | GATLYSPATRPSLADILKQSAR | 59 | | |
| <hr/> | | | | | |
| RESULT 14 | | | | | |
| ID | Q9V7W8 | PRELIMINARY; | PRT; | 460 AA. | |
| AC | Q9V7W8; | | | | |
| DT | 01-MAY-2000 (TREMBlrel. 13, | Created) | | | |
| DT | 01-MAY-2000 (TREMBlrel. 13, | Last sequence update) | | | |
| DT | 01-JUN-2001 (TREMBlrel. 17, | Last annotation update) | | | |
| DE | CG6796 PROTEIN. | | | | |
| GN | CG6796. | | | | |
| OS | Drosophila melanogaster (Fruit fly). | | | | |
| OC | Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; | | | | |
| OC | Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; | | | | |
| OC | Ephydroidea; Drosophilidae; Drosophila. | | | | |
| OX | NCBI_TaxID=7227; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | STRAIN-BERKELEY | | | | |
| RX | MEDLINE=20196006; PubMed=10731132; | | | | |
| RA | Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D., | | | | |
| RA | Amanatidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., | | | | |
| RA | George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., | | | | |
| RA | Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., | | | | |
| RA | Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., | | | | |
| RA | Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., | | | | |
| RA | April J.F., Aghayani A., An H.-J., Andrews-Franckoch C., Baldwin D., | | | | |
| RA | Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., | | | | |
| RA | Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S., | | | | |
| RA | Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P., | | | | |
| RA | Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., | | | | |
| RA | Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., | | | | |
| RA | de Fabios B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., | | | | |
| RA | Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., | | | | |
| RA | Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., | | | | |
| RA | Foster C., Gabrielian A.E., Garq N.S., Gelbart W.M., Glasser K., | | | | |
| RA | Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., | | | | |
| RA | Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., | | | | |
| RA | Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., | | | | |
| RA | Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., | | | | |
| RA | Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., | | | | |
| RA | Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., | | | | |
| RA | Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., | | | | |
| RA | Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., | | | | |
| RA | Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., | | | | |
| RA | Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., | | | | |
| RA | Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., | | | | |
| RA | Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., | | | | |
| RA | Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., | | | | |
| RA | Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., | | | | |
| RA | Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., | | | | |
| RA | Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J., | | | | |
| RA | Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., | | | | |
| RA | Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., | | | | |
| RA | Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., | | | | |
| RA | Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; | | | | |
| RT | "The genome sequence of Drosophila melanogaster". | | | | |
| SC | Science 287:2185-2195(2000). | | | | |
| DR | EMBL: AF003804; AAF57921.1; -- | | | | |
| DR | Flybase: Fgn0034177; CG6796. | | | | |
| DR | InterPro: IPR002106; AA_crna_ligase.II. | | | | |
| DR | InterPro: IPR002309; tRNA-synt.2. | | | | |
| DR | InterPro: IPR002312; tRNA-synt.asp. | | | | |
| <hr/> | | | | | |
| Db | 519 | TYISRLYPHLVKDSIQ | 535 | | |
| <hr/> | | | | | |
| Query Match 34.0%; Score 48; DB 12; Length 141; | | | | | |
| Best Local Similarity 47.4%; Pred. No. 18; | | | | | |
| Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps | | | | | |
| <hr/> | | | | | |
| QY | 7 | SRVTHPHLPKDIVRSIAK | 25 | | |
| | | : : | | | |
| DB | 64 | SKYWPNLRKDVKVSIRQC | 82 | | |
| <hr/> | | | | | |
| RESULT 13 | | | | | |
| ID | Q9KY27 | PRELIMINARY; | PRT; | 388 AA. | |
| AC | Q9KY27; | | | | |
| DT | 01-OCT-2000 (TREMBlrel. 15, | Created) | | | |
| DT | 01-OCT-2000 (TREMBlrel. 15, | Last sequence update) | | | |
| DT | 01-OCT-2000 (TREMBlrel. 15, | Last annotation update) | | | |
| | | | | | |

DR Pfam; PF00152; tRNA-synt_2; 1.
 DR Pfam; PF01336; tRNA_anti; 1.
 DR PRINTS; PR01042; TRNASYNTHASP.
 DR PROSITE; PS00179; AA_TRNA_LIGASE_IL1; UNKNOWN1.
 SQ SEQUENCE 460 AA; 52263 MW; CE0D8684BFBC90C CRC64;

Query Match 34.0%; Score 48; DB 5; Length 460;
 Best Local Similarity 77.8%; Pred. No. 61;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 10 THPHLPKDI 18
 :|||||:
 DB 403 SHPHLPKDL 411

RESULT 15
 Q9BIC9 PRELIMINARY; PRT; 421 AA.
 AC Q9BIC9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE SERINE PROTEINASE.
 OS Trichinella spiralis.
 OC Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
 OC Trichinellidae; Trichinella.
 OX NCBI_TaxID=6334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ISS413;
 RA Nagano I., Wu Z., Takahashi Y.;
 RT "Molecular expression and characterization of serine proteinase like
 protein from Trichinella spiralis."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY028974; AAK31787.1; -.
 SQ SEQUENCE 421 AA; 48035 MW; E1100A7BABE0091E CRC64;

Query Match 33.78%; Score 47.5; DB 5; Length 421;
 Best Local Similarity 45.0%; Pred. No. 66;
 Matches 9; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

QY 1 CGETYYS-RVTHPHLPKDIV 19
 ||| :||| |
 DB 30 CGNPYEPYLTNPYPNQIV 49

Search completed: March 4, 2002, 13:08:11
 Job time: 693 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:56:34 ; Search time 38.04 Seconds
(without alignments)
14.789 Million cell updates/sec

Title: US-09-701-623C-6

Perfect score: 141

Sequence: 1 CGETYSRVTHPLPKDIVRSTAKC 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 118 | 83.7 | 426 | 1 | US-08-336-583-2 |
| 2 | 118 | 83.7 | 426 | 5 | PCT-US95-13795-2 |
| 3 | 108 | 76.6 | 25 | 3 | US-09-100-414B-95 |
| 4 | 108 | 76.6 | 25 | 4 | US-09-303-323-95 |
| 5 | 108 | 76.6 | 42 | 3 | US-09-100-414B-98 |
| 6 | 108 | 76.6 | 42 | 3 | US-09-100-414B-99 |
| 7 | 108 | 76.6 | 42 | 3 | US-09-100-414B-100 |
| 8 | 108 | 76.6 | 42 | 4 | US-09-303-323-98 |
| 9 | 108 | 76.6 | 42 | 4 | US-09-303-323-99 |
| 10 | 108 | 76.6 | 42 | 4 | US-09-303-323-100 |
| 11 | 108 | 76.6 | 45 | 3 | US-09-100-414B-101 |
| 12 | 108 | 76.6 | 45 | 4 | US-09-303-323-101 |
| 13 | 108 | 76.6 | 46 | 3 | US-09-100-414B-96 |
| 14 | 108 | 76.6 | 46 | 3 | US-09-100-414B-97 |
| 15 | 108 | 76.6 | 46 | 4 | US-09-303-323-96 |
| 16 | 108 | 76.6 | 46 | 4 | US-09-303-323-97 |
| 17 | 108 | 76.6 | 63 | 3 | US-09-100-414B-102 |
| 18 | 108 | 76.6 | 63 | 4 | US-09-303-323-102 |
| 19 | 85 | 60.3 | 106 | 2 | US-08-232-539D-54 |
| 20 | 85 | 60.3 | 113 | 2 | US-08-232-539D-56 |
| 21 | 74.5 | 52.8 | 119 | 2 | US-08-464-025A-1 |
| 22 | 72 | 51.1 | 22 | 2 | US-08-232-539D-19 |
| 23 | 72 | 51.1 | 24 | 2 | US-08-232-539D-20 |
| 24 | 72 | 51.1 | 56 | 2 | US-08-232-539D-18 |
| 25 | 57 | 40.4 | 118 | 3 | US-08-466-151-1 |
| 26 | 54 | 38.3 | 1313 | 2 | US-08-244-537-2 |
| 27 | 54 | 38.3 | 2311 | 4 | US-08-934-386-9 |

| | | | | | | |
|----|------|------|------|---|-------------------|--------------------|
| 28 | 54 | 38.3 | 2325 | 3 | US-08-417-089-6 | Sequence 6, Appli |
| 29 | 54 | 38.3 | 2325 | 4 | US-08-695-651-6 | Sequence 6, Appli |
| 30 | 54 | 38.3 | 2325 | 4 | US-08-930-285-6 | Sequence 6, Appli |
| 31 | 54 | 38.3 | 2325 | 4 | US-08-695-421-6 | Sequence 6, Appli |
| 32 | 46 | 32.6 | 331 | 2 | US-08-646-981-17 | Sequence 17, Appli |
| 33 | 45 | 32.6 | 1205 | 4 | US-09-330-330-1 | Sequence 1, Appli |
| 34 | 45 | 31.9 | 561 | 3 | US-09-192-545-2 | Sequence 2, Appli |
| 35 | 45 | 31.9 | 2254 | 2 | US-08-677-010-3 | Sequence 3, Appli |
| 36 | 45 | 31.9 | 2254 | 2 | US-08-790-519-3 | Sequence 3, Appli |
| 37 | 44.5 | 31.6 | 1155 | 1 | US-08-094-948A-29 | Sequence 29, Appli |
| 38 | 44.5 | 31.6 | 1155 | 5 | PCT-US96-09319-29 | Sequence 29, Appli |
| 39 | 44 | 31.2 | 109 | 2 | US-08-646-981-6 | Sequence 6, Appli |
| 40 | 43.5 | 30.9 | 43 | 2 | US-08-488-161-31 | Sequence 31, Appli |
| 41 | 43.5 | 30.9 | 43 | 3 | US-09-273-685-31 | Sequence 31, Appli |
| 42 | 43.5 | 30.9 | 43 | 5 | PCT-US95-11934-31 | Sequence 31, Appli |
| 43 | 43 | 30.5 | 106 | 1 | US-08-399-106A-7 | Sequence 7, Appli |
| 44 | 43 | 30.5 | 106 | 1 | US-08-433-105A-7 | Sequence 7, Appli |
| 45 | 43 | 30.5 | 106 | 2 | US-08-434-869A-7 | Sequence 7, Appli |

ALIGNMENTS

RESULT 1
US-08-336-583-2
; Sequence 2, Application US/08336583
; Patent No. 5629415
; GENERAL INFORMATION:
; APPLICANT: HOLLIS, GREGORY F.
; APPLICANT: PATEL, MAYUR D.
; TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULIN E
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHRISTINE E. CARTY
; STREET: 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336.583
; FILING DATE: 09-NOV-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: CARTY, CHRISTINE E.
; REGISTRATION NUMBER: 36,099
; REFERENCE/DOCKET NUMBER: 19211
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-6734
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 426 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-336-583-2

Query Match 83.7%; Score 118; DB 1; Length 426;
Best Local Similarity 95.7%; Pred. No. 2.8e-11;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 GETYYSRVTHPLPKDIVRSTAK 24
Db 289 GETYYSRVTHPLPKDIVRSTAK 311

RESULT 2
PCT-US95-13795-2
; Sequence 2, Application PC/TUS9513795
; GENERAL INFORMATION:
; APPLICANT: HOLLIS, GREGORY F.
; APPLICANT: PATEL, MAYUR D.
; TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHRISTINE E. CARTY
; STREET: 126 E. LINCOLN AVENUE; P.O. BOX 2000
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13795
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CARTY, CHRISTINE E.
; REGISTRATION NUMBER: 36,099
; REFERENCE/DOCKET NUMBER: 19211Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-6734
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 426 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-13795-2

Query Match 83.7%; Score 118; DB 5; Length 426;
Best Local Similarity 95.7%; Pred. No. 2.8e-11;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GETYYSRVTHPLPKDIVRSIAK 24
||||| ||||||| ||||||| |||
Db 289 GETYYCRVTHPLPKDIVRSIAK 311

RESULT 3
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/100.414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-95

Query Match 76.6%; Score 108; DB 3; Length 25;
Best Local Similarity 72.0%; Pred. No. 4.1e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGETYYSRVTHPLPKDIVRSIAK 25
||||| ||||||| ||| |||
Db 1 CGETYQSRVTHPLPLMRSTTKC 25

RESULT 4
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303.323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100.414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 76.6%; Score 108; DB 4; Length 25;

Best Local Similarity 72.0%; Pred. No. 4.1e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPHLPKDIVRSIAKC 25
||||| ||||||| :||| ||
Db 1 CGETYQSRVTHPHLPALMRSTTKC 25

RESULT 5

US-09-100-414B-98
; Sequence 98, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-98

Query Match 76.6%; Score 108; DB 3; Length 42;
Best Local Similarity 72.0%; Pred. No. 7.6e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPHLPKDIVRSIAKC 25
||||| ||||||| :||| ||
Db 18 CGETYQSRVTHPHLPALMRSTTKC 42

RESULT 6

US-09-100-414B-99
; Sequence 99, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA

ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-99

Query Match 76.6%; Score 108; DB 3; Length 42;
Best Local Similarity 72.0%; Pred. No. 7.6e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYSRVTHPHLPKDIVRSIAKC 25
||||| ||||||| :||| ||
Db 18 CGETYQSRVTHPHLPALMRSTTKC 42

RESULT 7

US-09-100-414B-100
; Sequence 100, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-09-100-414B-100

Query Match 76.6%; Score 108; DB 3; Length 42;
Best Local Similarity 72.0%; Pred. No. 7.6e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||| ||||| :||| ||
Db 18 CGETYQSRVTHPLPRALMRSTTKC 42

RESULT 8

US-09-303-323-98
; Sequence 98, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-303-323-98

Query Match 76.6%; Score 108; DB 4; Length 42;
Best Local Similarity 72.0%; Pred. No. 7.6e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||| ||||| :||| ||
Db 18 CGETYQSRVTHPLPRALMRSTTKC 42

RESULT 9

US-09-303-323-99
; Sequence 99, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998

NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-99

Query Match 76.6%; Score 108; DB 4; Length 42;
Best Local Similarity 72.0%; Pred. No. 7.6e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||| ||||| :||| ||
Db 18 CGETYQSRVTHPLPRALMRSTTKC 42

RESULT 10

US-09-303-323-100
; Sequence 100, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998

ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-101

Query Match 76.6%; Score 108; DB 4; Length 42;
Best Local Similarity 72.0%; Pred. No. 7.6e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||||| :||| ||
Db 18 CGETYQSRVTHPLPRALMRSTTKC 42

RESULT 11
US-09-100-414B-101
Sequence 101, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
REFERENCE/DOCKET NUMBER: 1151-4157
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-751-6849
TELEFAX: 212-758-4800
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-101

Query Match 76.6%; Score 108; DB 3; Length 45;
Best Local Similarity 72.0%; Pred. No. 8.2e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||||| :||| ||
Db 21 CGETYQSRVTHPLPRALMRSTTKC 45

RESULT 12
US-09-303-323-101
Sequence 101, Application US/09303323
Patent No. 6228987
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
REFERENCE/DOCKET NUMBER: 1151-4157
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-101

Query Match 76.6%; Score 108; DB 4; Length 45;
Best Local Similarity 72.0%; Pred. No. 8.2e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||||| :||| ||
Db 21 CGETYQSRVTHPLPRALMRSTTKC 45

RESULT 13
US-09-100-414B-96
Sequence 96, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
REFERENCE/DOCKET NUMBER: 1151-4157
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-96

Query Match 76.6%; Score 108; DB 3; Length 46;
Best Local Similarity 72.0%; Pred. No. 8.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||| :||| ||
Db 22 CGETYQSRVTHPLPLMRSTTKC 46

RESULT 14
US-09-100-414B-97
Sequence 97, Application US/09100414B
Patent No. 6025458
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-97

Query Match 76.6%; Score 108; DB 3; Length 46;
Best Local Similarity 72.0%; Pred. No. 8.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||| :||| ||
Db 22 CGETYQSRVTHPLPLMRSTTKC 46

RESULT 15
US-09-303-323-96
Sequence 96, Application US/09303323
Patent No. 6228987
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-96

Query Match 76.6%; Score 108; DB 4; Length 46;
Best Local Similarity 72.0%; Pred. No. 8.5e-11;
Matches 18; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 CGETYYSRVTHPLPKDIVRSIAKC 25
||||| ||||| :||| ||
Db 22 CGETYQSRVTHPLPLMRSTTKC 46

Search completed: March 4, 2002, 12:56:35
Job time: 257 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:58:20 ; Search time 42.32 Seconds
(without alignments)
44.999 Million cell updates/sec

Title: US-09-701-623c-7

Perfect score: 145

Sequence: 1 CGEGYQSRVDHPFKPIVRSITK 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_68.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|----------|------------------------|
| 1 | 122 | 84.1 | 429 | 1 EHRT | Ig epsilon chain C |
| 2 | 98 | 67.6 | 107 | 2 I68730 | IgE chain C3 regio |
| 3 | 98 | 67.6 | 107 | 2 I68726 | IgE chain C3 regio |
| 4 | 98 | 67.6 | 388 | 1 EHMS | Ig epsilon chain C |
| 5 | 98 | 67.6 | 388 | 2 S38864 | Ig epsilon chain C |
| 6 | 85 | 58.6 | 423 | 1 EHMS | Ig epsilon chain C |
| 7 | 83 | 57.2 | 426 | 2 I36948 | Ig epsilon chain C |
| 8 | 80 | 55.2 | 428 | 1 EHHU | Ig epsilon chain C |
| 9 | 51 | 35.2 | 245 | 2 I55951 | MHC class II E-bet |
| 10 | 51 | 35.2 | 264 | 2 A60497 | H-2 class II histocomp |
| 11 | 51 | 35.2 | 284 | 2 SI0989 | class II histocomp |
| 12 | 49 | 33.8 | 345 | 1 HLCHB4 | MHC class I histoc |
| 13 | 49 | 33.8 | 355 | 2 I51309 | major histocompati |
| 14 | 49 | 33.8 | 355 | 2 T28149 | MHC class I histoc |
| 15 | 49 | 33.8 | 355 | 2 T28152 | MHC class I histoc |
| 16 | 49 | 33.8 | 549 | 2 S04845 | Ig heavy chain prote |
| 17 | 49 | 33.8 | 1208 | 2 T23222 | hypothetical prote |
| 18 | 48.5 | 33.4 | 567 | 2 T33400 | protein kinase C h |
| 19 | 48.5 | 33.4 | 597 | 2 T33399 | protein kinase C h |
| 20 | 48.5 | 33.4 | 704 | 1 S60117 | protein kinase C (|
| 21 | 48 | 33.1 | 93 | 2 I54421 | MHC RT1.B-beta2 - |
| 22 | 48 | 33.1 | 110 | 2 S43147 | Ig epsilon chain - |
| 23 | 48 | 33.1 | 152 | 2 SI4236 | Ig gamma-1 chain C |
| 24 | 48 | 33.1 | 237 | 2 C60497 | H-2 class II histoc |
| 25 | 48 | 33.1 | 266 | 2 A39260 | MHC class II histoc |
| 26 | 48 | 33.1 | 266 | 2 B39260 | MHC class II histoc |
| 27 | 48 | 33.1 | 324 | 1 GLMS | Ig gamma-1 chain C |
| 28 | 48 | 33.1 | 393 | 1 GLMSM | Ig gamma-1 chain C |
| 29 | 48 | 33.1 | 444 | 2 PC4436 | monoclonal antibod |

30 47.5 32.8 1198 2 T49726 hypothetical prote
31 47 32.4 322 2 PS0019 Ig gamma-2a chain
32 47 32.4 322 2 T26467 hypothetical prote
33 47 32.4 2254 2 D86215 protein T6D22.14 (

ALIGNMENTS

RESULT 1

EHRT

Ig epsilon chain C region - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 16-Jul-1999

C:Accession: A93442; A90937; A02143

R:Hellman, L.; Pettersson, U.; Engstrom, A.; Karlsson, T.; Bennich, H.

Nucleic Acids Res. 10, 6041-6049, 1982

A:Title: Structure and evolution of the heavy chain from rat immunoglobulin E.

A:Reference number: A93442; MUID:83064537

A:Accession: A93442

A:Molecule type: mRNA

A:Residues: 1-429 <HEL>

A:Experimental source: Strain LOU/c/Wsl, immunocytoma IR2

R:Kindsvogel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.

DNA 1, 335-343, 1982

A:Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction

A:Reference number: A90937; MUID:83182019

A:Contents: myeloma IR162

A:Accession: A90937

A:Molecule type: mRNA

A:Residues: 'N',169-307,'L',309-342 <KIN>

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (

chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into

C:Superfamily: immunoglobulin C region: immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F:19-80/Domain: immunoglobulin homology <IM1>

F:118-186/Domain: immunoglobulin homology <IM2>

F:223-291/Domain: immunoglobulin homology <IM3>

F:327-398/Domain: immunoglobulin homology <IM4>

F:46,99,170,240,265,369,419/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match

Best Local Similarity 84.1%; Score 122; DB 1; Length 429;

Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPFKPIVRSITK 24

||||| |||||||||

Db 284 GEGYQSRVDHPFKPIVRSITK 306

RESULT 2

I68730

IgE chain C3 region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000

C:Accession: I68730

R:Shinkai, Y.; Nakauchi, H.; Honjo, T.; Okumura, K.

Immunogenetics 27, 288-292, 1988

A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic aci

A:Reference number: I54443; MUID:88152907

A:Residues: 1-358, 'L', 360-428 <MAX>
A:Cross-references: GB:J00222; NID:gl84755
A:Note: this sequence difference may be due to polymorphism
R:Bennich, H.H.; Johansson, S.G.O.; Von Bahr-Lindstrom, H.
in Immediate Hypersensitivity: Modern Concepts and Developments, Bach, M.K., ed., pp.
A:Reference number: A94418
A:Accession: A94418
A:Molecule type: Protein
A:Residues: 'GAWTL', 6, 'X', 8-16, 'B', 18-43, 'B', 45-52, 55-92, 95-97, 'B', 99-121, 'B', 123, 'L',
A:Experimental source: myeloma protein Nd
R:Kenten, J.H.; Molgaard, H.V.; Houghton, M.; Derbyshire, R.B.; Viney, J.; Bell, L.O.
Proc. Natl. Acad. Sci. U.S.A. 79, 6661-6665, 1982
A:Title: Cloning and sequence determination of the gene for the human immunoglobulin
A:Reference number: A93933; MUID:83065234
A:Accession: B93933
A:Molecule type: mRNA
A:Residues: 1-40; 68-114; 427-428 <KEN>
A:Cross-references: GB:L00022; NID:gl85035
R:Ikeyama, S.
FEBS Lett. 224, 306-310, 1987
A:Title: Purification and characterization of a recombinant human IgE Fc-epsilon frag
A:Reference number: S02438; MUID:88083554
A:Accession: S02438
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 98-352 <IKE>
R:Zhang, K.; Max, E.E.; Cheah, H.K.; Saxon, A.
J. Biol. Chem. 269, 456-462, 1994
A:Title: Complex alternative RNA splicing of epsilon-immunoglobulin transcripts produ
A:Reference number: A53116; MUID:94103254
A:Accession: A53116
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 320-428 <H2>
A:Experimental source: myeloma U266-derived cell line AF-10
A:Note: sequence extracted from NCBI backbone (NCBI:141701, NCBI:141702)
R:Hellman, L.
Eur. J. Immunol. 23, 159-167, 1993
A:Title: Characterization of four novel epsilon chain mRNA and a comparative analysis
A:Reference number: A46536; MUID:93122085
A:Accession: C46536
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 382-426 <HEL>
A:Cross-references: GB:S55273; NID:g263166; PIDN:AAB24857.1; PID:g263167
A:Experimental source: B cell myeloma U-266
A:Note: sequence extracted from NCBI backbone (NCBI:125297)
A:Accession: D46536
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 382-391 <HE2>
A:Cross-references: GB:S55276; PIDN:AAB24858.1; PID:g263169
A:Experimental source: B cell myeloma U-266
A:Note: sequence extracted from NCBI backbone (NCBI:125299)
A:Accession: A46536
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 401-428 <HE3>
A:Cross-references: GB:S53497; NID:g263162; PIDN:AAB24855.1; PID:g263163
A:Experimental source: B cell myeloma U-266
A:Note: sequence extracted from NCBI backbone (NCBI:123483)
C:Genetics:
A:Gene: GDB:IGHE
A:Cross-references: GDB:119335; OMIM:147180
A:Map position: 14q32.33-14q32.33
A:Introns: 1/1; 104/1; 211/1; 319/1
C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (C
chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterotrimer; immunogl
F; 22-87/Domain: immunoglobulin homology <IM1>
F; 128-195/Domain: immunoglobulin homology <IM2>
F; 232-301/Domain: immunoglobulin homology <IM3>

```
F;338-407/Domain: immunoglobulin homology <IM4>
F;14/Disulfide bonds: Interchain (to light chain) #status predicted
F;15-105,29-88,135-193,239-299,345-405/Disulfide bonds: #status predicted
F;21,49,99,146,252,275/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;121,209/Disulfide bonds: Interchain (to heavy chain) #status predicted
```

Query Match 55.2%; Score 80; DB 1; Length 428;
Best Local Similarity 60.9%; Pred. No. 0.00046;
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHFPKPIVRSITK 24
 |||||
DG 294 GETYQCQRVTTHPLPRALMRSTTK 316

RESULTS

MHC class II E-beta protein - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 21-Jan-2000
 C:Accession: I55951
 R:Robertson, K.A.; McMaster, W.R.
 J. Immunol. 135, 4095-4099, 1985
 A:Title: Complete structure of a rat Rtl E-beta chain: Extensive conservation of MHC class II E-beta chain
 A:Reference number: I55951; MUID:86060895
 A:Accession: I55951
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: mRNA
 A:Residues: 1-245 <RES>
 A:Cross-references: GB:M12382; NID:g205437; PIDN:AAA41610.1; PID:g205438
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology
 F:118-183/Domain: immunoglobulin homology <IMM>

Query Match 35.2%; Score 51; DB 2; Length 245;
Best Local Similarity 47.1%; Pred. No. 4.6;
Matches 8; Conservative 3; Mismatches 6; Indels

QY 2 GEGYQSRVDHHPFKPI 18
 ||| :||| | | :
 Db 176 GEVYTCOVEHPSPV 192

RESULT

A60497
H-2 class II histocompatibility antigen RT1.D-mu beta chain precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jan-2000
C:Accession: A60497
R:Holowachuk, E.W.; Greer, M.K.
Diabetes 38, 267-271, 1989
A:Title: Unaltered class II histocompatibility antigens and pathogenesis of IDDM in BB
A:Reference number: A60497; MUID:89121214
A:Accession: A60497
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-264 <HOL>
C:Superfamily: Class II histocompatibility antigen; Immunoglobulin homology
F:137-202/Domain: Immunoglobulin homology <IIMM>

Query Match 35.2%; Score 51; DB 2; Length 264;
Best Local Similarity 47.1%; Pred. No. 5;
Matches 8; Conservative 3; Mismatches 6; Indels

QY 2 GEGYQSRVDHPHPKPI 18
||| :||| | | :
Db 195 GEVYTCOVHPSPV 211

RESULT 17

RESOLUTION
11
S10989

class II histocompatibility antigen RT1-D beta-I chain precursor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 21-Jan-2000
C/Accession: S10989
R/Syha-Jedelhauser, J.; Reske, K.
Nucleic Acids Res. 18, 4598, 1990
A/Title: Sequence of rat cDNA clone pLR-beta-112 coding for the RT1.D-beta' chain.
A/Reference number: S10989.MUID:90356406

A: Residues: 1-264 <SVH>
A: CROSS-references: EMBL:X53054; NID:q57169; PIDN:CAA37221.1; PID:q57170
C: Superfamily: class II histocompatibility antigen; immunoglobulin homology
C: keywords: glycoprotein; heterodimer; transmembrane protein
F: 1-264/Domain: signal sequence #status: predicted <SIG>
F: 17-264/Product: class II histocompatibility antigen, RFL-D beta-I chain #status: pre
F: 137-202/Domain: immunoglobulin homology <IMM>
F: 229-246/Domain: transmembrane #status: predicted <TMW>
F: 45/Binding site: carboxydrate (Asn) (covalent) #status: predicted

| | | | | |
|-----------------------|--------|--------------|-------|-------------|
| Query Match | 35.2% | Score 51; | DB 2; | Length 264; |
| Best Local Similarity | 47.1%; | Pred. No. 5; | | |
| Matches | 8; | Conservative | 3; | Mismatches |
| | | | 6; | Indels |
| | | | 0; | Gaps |

QY 2 GEGYQSRVDHPHFPI 18
 ||| :||| | | :
 db 195 GEVYTCOVEHPSPV 211

RESULT 12

HI.CHR4
RESOL1

MHC class I histocompatibility antigen B-F IV alpha chain precursor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-Jun-1999
C:Accession: A45846: S01172
R:Kroemer, G.; Zoorob, R.; Auffray, C.
Immunogenetics 31, 405-409, 1990
A:Title: Structure and expression of a chicken MHC class I gene.
A:Reference number: A45846; MUID:90316612

A: MOLECULE TYPE: DNA
A: RESIDUES: 1-345 <KPO>

A; residues: I-343 <KRO>
A:Cross-references: GB:M31012

R:Guillemot, F.: Billault, A.

EMBO J. 7, 2775-2785, 1988

A>Title: A molecular map of the chicken m

A; Reference number:

A; Accession: S01172

A; Molecule type: mRNA

A;Residues: 1-345 <GUI>

A;Cross-references: EMBL:X12780; NID:g63089; PIDN:CAA31272.1; PID:g63090

C;Genetics:

A:Map position: 16
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: glycoprotein; heterodimer; transmembrane protein
F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-345/Product: class I histocompatibility antigen B-F IV alpha chain #status predicted
F:23-301/Domain: extracellular #status predicted <EXT>
F:23-110/Domain: alpha-1 <EXT>
F:111-201/Domain: alpha-2 <EXT>
F:214-279/Domain: immunoglobulin homology <IMM>
F:302-324/Domain: transmembrane #status predicted <TM>
F:325-345/Domain: intracellular #status predicted <INT>
F:59,107/Binding site: carbohydrate (ASN) (covalent) #status predicted
F:121-183,221-277/disulfide bonds: #status predicted

Query Match 33.8%; Score 49; DB 1; Length 345;
Best Local Similarity 50.0%; Pred. No. 13;
Matches 8; Conservative 3; Mismatches 5; Indels

QY 2 GEGYQSRVDHPHF KP 17

Db 272 GDKYQCRVEHASLPQP 287

! : || || : | | : |

RESULT 13

T51309

major histocompatibility complex class I glycoprotein haplotype B21 - chicken

C:Species: Gallus gallus (chicken)

C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 21-Jan-2000

C:Accession: I51309

R:Fulton, J.E.; Thacker, E.L.; Bacon, L.D.; Hunt, H.D.

Eur. J. Immunol. 25, 2069-2076, 1995

A:Title: Functional analysis of avian class I (BFIV) glycoproteins by epitope tagging and

A:Reference number: I51309; MUID:95347411

A:Accession: I51309

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-355 <FUI>

A:Cross-references: GB:S78682; NID:gl042200; PIDN:AAB34945.1; PID:gl042201

C:Genetics:

A:Gene: BFIV21

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

F:213-278/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 33.8%; Score 49; DB 2; Length 355;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKP 17

! : || || : | | : |

Db 271 GDKYQCRVEHASLPQP 286

RESULT 14

T28149

MHC class I histocompatibility antigen B-F alpha chain 2 - chicken

C:Species: Gallus gallus (chicken)

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000

C:Accession: T28149

R:Milne, S.; Kaufman, J.; Beck, S.

submitted to the EMBL Data Library, May 1998

A:Description: DNA sequencing and analysis of the chicken major histocompatibility compl

A:Reference number: 220475

A:Accession: T28149

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-355 <MIL>

A:Cross-references: EMBL:AL023516; PIDN:CAA18969.1

A:Experimental source: clone cB12

C:Genetics:

A:Gene: Bfa2

A:Map position: 16

A:Introns: 22/1; 110/1; 201/1; 292/1; 328/1; 339/1; 350/1

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 33.8%; Score 49; DB 2; Length 355;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKP 17

! : || || : | | : |

Db 271 GDKYQCRVEHASLPQP 286

RESULT 15

T28152

MHC class I histocompatibility antigen B-F IV alpha chain precursor - chicken

C:Species: Gallus gallus (chicken)

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000

C:Accession: T28152

R:Milne, S.; Kaufman, J.; Beck, S.

submitted to the EMBL Data Library, May 1998
A:Description: DNA sequencing and analysis of the chicken major histocompatibility co
A:Reference number: 220475

A:Accession: T28152

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-355 <MIL>

A:Cross-references: EMBL:AL023516; PIDN:CAA18972.1

A:Experimental source: clone cB12

C:Genetics:

A:Gene: Bfa1

A:Map position: 16

A:Introns: 22/1; 110/1; 201/1; 292/1; 328/1; 339/1; 350/1

C:Superfamily: class I histocompatibility antigen; immunoglobulin homology

Query Match

Best Local Similarity 33.8%; Score 49; DB 2; Length 355;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKP 17

! : || || : | | : |

Db 271 GDKYQCRVEHASLPQP 286

Search completed: March 4, 2002, 12:58:21

Job time: 248 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:08:42 ; Search time 24.65 Seconds
(without alignments)
37.185 Million cell updates/sec

Title: US-09-701-623C-7

Perfect score: 145
Sequence: 1 CGEGYOSRVDHPHPPKIVRSITKC 25

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 122 | 84.1 | 429 | 1 EPC_RAT | P01855 rattus norv |
| 2 | 98 | 67.6 | 421 | 1 EPC_MOUSE | P06336 mus musculu |
| 3 | 80 | 55.2 | 428 | 1 EPC_HUMAN | P01854 homo sapien |
| 4 | 51 | 35.2 | 264 | 1 HB2D_RAT | P18211 rattus norv |
| 5 | 49 | 33.8 | 345 | 1 HALF_CHICK | P15979 gallus gall |
| 6 | 48.5 | 33.4 | 704 | 1 KPC1_CAEEL | P34722 caenorhabdi |
| 7 | 48 | 33.1 | 324 | 1 GC1_MOUSE | P01868 mus musculu |
| 8 | 48 | 33.1 | 393 | 1 GC1M_MOUSE | P01869 mus musculu |
| 9 | 47 | 32.4 | 322 | 1 GCA_RAT | P20760 rattus norv |
| 10 | 46.5 | 32.1 | 345 | 1 DER2_HUMAN | Q9h3k2 homo sapien |
| 11 | 46 | 31.7 | 116 | 1 SOR_METJA | Q58151 methanococc |
| 12 | 46 | 31.7 | 277 | 1 VNST_CVBO | P18517 bovine coro |
| 13 | 46 | 31.7 | 336 | 1 GCB_MOUSE | P01866 mus musculu |
| 14 | 46 | 31.7 | 370 | 1 DCUP_SCHPO | Q9usj5 schizosacch |
| 15 | 46 | 31.7 | 373 | 1 GCB_MOUSE | Q02395 mus musculu |
| 16 | 46 | 31.7 | 405 | 1 GCB_MOUSE | P01867 mus musculu |
| 17 | 46 | 31.7 | 516 | 1 MEPA_XENLA | Q03414 xenopus lae |
| 18 | 46 | 31.7 | 593 | 1 MTF2_HUMAN | Q9y483 homo sapien |
| 19 | 45 | 31.0 | 116 | 1 YK14_YEAST | P36078 saccharomyc |
| 20 | 45 | 31.0 | 330 | 1 GCAA_MOUSE | P01863 mus musculu |
| 21 | 45 | 31.0 | 399 | 1 GCAM_MOUSE | P01865 mus musculu |
| 22 | 44.5 | 30.7 | 387 | 1 GCP_RICPR | Q92ea8 rickettsia |
| 23 | 44 | 30.3 | 104 | 1 Y4EB_RHISN | P55425 rhizobium s |
| 24 | 44 | 30.3 | 263 | 1 HB2B_RAT | P29826 rattus norv |
| 25 | 44 | 30.3 | 275 | 1 Y760_METJA | Q58170 methanococc |
| 26 | 44 | 30.3 | 290 | 1 GC3_HUMAN | P01860 homo sapien |
| 27 | 44 | 30.3 | 326 | 1 GC2_HUMAN | P01859 homo sapien |
| 28 | 44 | 30.3 | 329 | 1 GC3_MOUSE | P22436 mus musculu |
| 29 | 44 | 30.3 | 330 | 1 GC1_HUMAN | P01857 homo sapien |
| 30 | 44 | 30.3 | 371 | 1 HA12_RAT | P16391 rattus norv |
| 31 | 44 | 30.3 | 371 | 1 RIBD_BACAM | P70814 b riboflavi |
| 32 | 44 | 30.3 | 384 | 1 GBA2_PEA | O04279 pisum sativ |
| 33 | 44 | 30.3 | 398 | 1 GC3M_MOUSE | P03987 mus musculu |

RESULT 1

| ID | EPC_RAT | STANDARD; | PRT; | 429 AA. |
|----|----------------------------------------------------------------------------|-----------|------|---------|
| DT | 21-JUL-1986 (Rel. 01, Created) | | | |
| DT | 21-JUL-1986 (Rel. 01, Last sequence update) | | | |
| DT | 15-JUL-1999 (Rel. 38, Last annotation update) | | | |
| DE | IG EPSILON CHAIN C REGION. | | | |
| OS | Rattus norvegicus (Rat). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. | | | |
| OX | NCBI_TaxID=10116; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2). | | | |
| RC | STRAIN=LOU/C/WSL; | | | |
| RX | MEDLINE=83064537; PubMed=6292865; | | | |
| RA | Heilmann L., Pettersson U., Engstroem A., Karlsson T., Bennich H.; | | | |
| RT | "Structure and evolution of the heavy chain from rat immunoglobulin | | | |
| RT | E."; | | | |
| RL | Nucleic Acids Res. 10:6041-6049(1982). | | | |
| RN | [2] | | | |
| RP | SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162). | | | |
| RX | MEDLINE=83182019; PubMed=6820340; | | | |
| RA | Kindsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.; | | | |
| RT | "A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: | | | |
| RT | construction, identification, and DNA sequence."; | | | |
| RL | DNA 1:335-343(1982). | | | |
| RN | [3] | | | |
| RP | SEQUENCE OF 205-306 FROM N.A. | | | |
| RX | MEDLINE=82174576; PubMed=6803238; | | | |
| RA | Heilmann L., Pettersson U., Bennich H.; | | | |
| RT | "Characterization and molecular cloning of the mRNA for the heavy | | | |
| RT | (epsilon) chain of rat immunoglobulin E."; | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 79:1264-1268(1982). | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (see http://www.isb-sib.ch/announce/ | | | |
| CC | or send an email to license@isb-sib.ch). | | | |
| CC | ----- | | | |
| CC | EMBL; J00744; AAA41379.1; ALT_INIT. | | | |
| DR | PIR; A02143; EHRT. | | | |
| DR | InterPro; IPR003006; Iq_MHC. | | | |
| DR | InterPro; IPR003597; Iq_cl. | | | |
| DR | InterPro; IPR003600; Iq_like. | | | |
| DR | Pfam; PF00047; Iq; 4. | | | |
| DR | SMART; SM00407; IGcl; 1. | | | |
| DR | SMART; SM00410; IG_like; 3. | | | |
| DR | PROSITE; PS00290; IG_MHC; 3. | | | |
| KW | Immunoglobulin domain; Immunoglobulin C region. | | | |
| FT | NON_TER 1 | | | |
| FT | CONFLICT 168 168 R -> N (IN REF. 2). | | | |
| FT | CONFLICT 308 308 P -> L (IN REF. 2). | | | |

P01875 gallus gall
Q25378 lytechinus
P53585 caenorhabdi
P81869 cavia porce
P17075 homo sapien
P23403 xenopus lae
P55828 drosophila
P06346 mus musculu
P01920 homo sapien
P06343 mus musculu
P06345 mus musculu
P06344 mus musculu

SQ SEQUENCE 429 AA; 48671 MW; D2970B34EF8A72B0 CRC64;

Query Match 84.1%; Score 122; DB 1; Length 429;

Best Local Similarity 95.7%; Pred. NO. 1.5e-10; Mismatches 1; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
||||| ||||||| ||||||| |||||||

Db 284 GEGYQSRVDHPHPKPIVRSITK 306

RESULT 2

```
EPC_MOUSE
ID EPC_MOUSE STANDARD; PRT; 421 AA.
AC P06336; P01856;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG EPSILON CHAIN C REGION.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10030;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=84236092; PubMed=6329728;
RA Ishida N., Ueda S., Hayashida H., Miyata T., Honjo T.;
RT "The nucleotide sequence of the mouse immunoglobulin epsilon gene:
RL comparison with the human epsilon gene sequence.";
RN EMBO J. 1:1117-1123(1982).
RN [2]
RP REVISIONS.
RA Honjo T.;
RL Submitted (APR-1986) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE OF 34-421 FROM N.A.
RX MEDLINE=8311774; PubMed=6818553;
RA Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;
RT "Cloning and nucleotide sequence of mouse immunoglobulin epsilon
chain cDNA.";
RN Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856(1982).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X01857; CAA25977.1; -;
DR EMBL; X01857; CAA25978.1; -;
DR PIR; A02145; EHMS.
DR PIR; A02144; EHMS.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig_4.
DR SMART; SM00407; Igcl; 2.
DR SMART; SM00410; Ig_like; 2.
DR PROSITE; PS00290; IG_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1 90
FT DOMAIN 1 90 CH1.
FT DOMAIN 91 197 CH2.
FT DOMAIN 198 304 CH3.
FT DOMAIN 305 421 CH4.
FT DISULFID 23 75 BY SIMILARITY.
FT DISULFID 121 180 BY SIMILARITY.
FT DISULFID 226 285 BY SIMILARITY.
FT DISULFID 330 392 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 365 365 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 421 AA; 47320 MW; 8F909E1F30A06B47 CRC64;
```

Query Match 67.6%; Score 98; DB 1; Length 421;
Best Local Similarity 82.6%; Pred. NO. 4.8e-07;
Matches 19; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
| ||| |||| ||||||| |||||

Db 280 GEGYQSRVDHPHPKPIVRSITK 302

RESULT 3

```
EPC_HUMAN
ID EPC_HUMAN STANDARD; PRT; 428 AA.
AC P01854;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE IG EPSILON CHAIN C REGION.
GN IGHE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=83168897; PubMed=6300763;
RA Seno M., Kurokawa T., Ono Y., Onda H., Sasada R., Igarashi K.,
RA Kikuchi M., Sugino Y., Nishida Y., Honjo T.;
RT "Molecular cloning and nucleotide sequencing of human immunoglobulin
epsilon chain cDNA.";
RN Nucleic Acids Res. 11:719-726(1983).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=83001945; PubMed=6288268;
RA Max E.E., Battey J., Ney R., Kirsch I.R., Leder P.;
RT "Duplication and deletion in the human immunoglobulin epsilon genes.";
RN Cell 29:691-699(1982).
RN [3]
SEQUENCE FROM N.A.
RX MEDLINE=84236029; PubMed=6234164;
RA Flanagan J.G., Rabbitts T.H.;
RT "The sequence of a human immunoglobulin epsilon heavy chain constant
region gene, and evidence for three non-allelic genes.";
RN EMBO J. 1:655-660(1982).
RN [4]
SEQUENCE FROM N.A.
RX MEDLINE=84207910; PubMed=6327276;
RA Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.;
RT "Long terminal repeat-like elements flank a human immunoglobulin
epsilon pseudogene that lacks introns.";
RN EMBO J. 1:1539-1544(1982).
RN [5]
PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND).
RA Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.;
RL (In) Bach M.K. (eds.);
RL Immediate hypersensitivity: modern concepts and developments, pp.1-36,
RL Marcel Dekker, New York (1978).
RN [6]
SEQUENCE OF 1-40; 68-114 AND 427-428 FROM N.A.
RX MEDLINE=83065234; PubMed=6815656;
RA Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J.,
RA Bell L.O., Gould H.J.;
RT "Cloning and sequence determination of the gene for the human
```


class II beta genes are closely linked to the class I genes and the nucleolar organizer.";
EMBO J. 7:2775-2785(1988).

[2]
SEQUENCE FROM N.A.
MEDLINE=90316612; PubMed=2370087;
RA Kroemer G., Zocrob R., Auffray C.;
RT "Structure and expression of a chicken MHC class I gene.";
RL Immunogenetics 31:405-409(1990).
CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO
CC THE IMMUNE SYSTEM.
CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
CC MICROGLOBULIN).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

EMBL: X12780; CAA31272.1; -;
DR EMBL: M31012; AAA48947.1; -;
DR PIR: S01172; H1CHB4.
DR PIR: A45846; A15846.
DR HSSP: P10318; LR0L.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR001039; MHC_I.
DR Pfam: PF00047; Ig; 1.
DR Pfam: PF00129; MHC_I; 1.
DR ProDom: PD000050; MHC_I; 1.
DR SMART: SM00407; IG_C1; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KW MHC I; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 22
FT CHAIN 23 345
FT FT
FT ALPHA CHAIN.
FT DOMAIN 23 110
FT EXTRACELLULAR ALPHA-1.
FT DOMAIN 111 201
FT EXTRACELLULAR ALPHA-2.
FT DOMAIN 202 292
FT EXTRACELLULAR ALPHA-3.
FT DOMAIN 293 301
FT CONNECTING PEPTIDE.
FT TRANSMEM 302 324
FT DOMAIN 325 345
FT CYTOPLASMIC.
FT DISULFID 121 183
FT DISULFID 221 277
FT BY SIMILARITY.
FT CARBOHYD 59 59
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 107 107
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 339 339 P -> PDREGGSSSSST (IN REF. 2).
SQ SEQUENCE 345 AA; 38246 MW; 430DCCF8091B69A4 CRC64;

Query Match 33.8%; Score 49; DB 1; Length 345;
Best Local Similarity 50.0%; Pred. No. 5.8;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVJHPFPKP 17
Db 272 GDYIQRVHSLPQP 287
1: ||||| :||
:

RESULT 6
KPC1_CAEEL STANDARD; PRT; 704 AA.
AC P34722;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROTEIN KINASE C-LIKE 1 (EC 2.7.1.-) (PKC).
GN TPA-1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditiidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=95387388; PubMed=7658466;
RA Sano T., Tabuse Y., Nishiwaki K., Miwa J.;
RT "The tpa-1 gene of Caenorhabditis elegans encodes two proteins
RT similar to Ca(2+)-independent protein kinase Cs: evidence by complete
RT genomic and complementary DNA sequences of the tpa-1 gene.";
RL J. Mol. Biol. 251:477-485(1995).
RN [2]
RN SEQUENCE OF 148-704 FROM N.A.
RX MEDLINE=89186920; PubMed=2538925;
RA Tabuse Y., Nishiwaki K., Miwa J.;
RT "Mutations in a protein kinase C homolog confer phorbol ester
RT resistance on Caenorhabditis elegans.";
RL Science 243:1713-1716(1989).
RN [3]
RP REVISIONS TO REF.2.
RA Miwa J.;
CC Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PKC IS ACTIVATED BY DIACYLGLYCEROL WHICH IN TURN
CC PHOSPHORYLATES A RANGE OF CELLULAR PROTEINS. PKC ALSO SERVES AS
CC THE RECEPTOR FOR PHORBOL ESTERS, A CLASS OF TUMOR PROMOTERS.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; TPA-1A (SHOWN HERE) AND TPA-1B;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC BINDING DOMAINS.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC PKC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

EMBL: D49525; BAA08470.1; -;
DR EMBL: D49525; BAA08471.1; -;
DR EMBL: D14815; BAA03556.1; -;
DR HSSP: P28867; 1PTR.
DR InterPro: IPR002219; DAG_PE-bind.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR000961; Pkinase_C.
DR Pfam: PF00130; DAG_PE-bind; 2.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00433; pkinase; 1.
DR PRINTS; PR00008; DAGPEDOMAIN.
DR SMART; SM00109; C1; 2.
DR SMART; SM00220; S-TKC; 1.
DR PROSITE; PS00133; S-TKC_X; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Zinc;
KW Phorbol-ester binding; Repeat; Alternative splicing; Phosphorylation.
FT DOMAIN 166 215 PHORBOL-ESTER AND DAG BINDING 1.
FT DOMAIN 238 287 PHORBOL-ESTER AND DAG BINDING 2.
FT DOMAIN 375 634 PROTEIN KINASE.
FT NP_BIND 381 389 ATP (BY SIMILARITY).
FT BINDING 404 404 ATP (BY SIMILARITY).
FT ACT_SITE 499 499 BY SIMILARITY.
FT MOD_RES 89 89 PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT MOD_RES 139 139 PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT MOD_RES 324 324 PHOSPHORYLATION (AUTO-) (POTENTIAL).
FT VARSPPLIC 1 137 MISSING (IN ISOFORM TPA-1B).
SQ SEQUENCE 704 AA; 80298 MW; C44F2E25F58057E3 CRC64;

Query Match 33.4%; Score 48.5; DB 1; Length 704;
 Best Local Similarity 37.9%; Pred. No. 14;
 Matches 11; Conservative 2; Mismatches 11; Indels 5; Gaps 1;

QY 2 GEGYQSRVD-----HPHFKPPIVRSITK 25

DB 577 GGEDELFDLSILNERYPFKTKISKAAC 605

RESULT 7

GCL_MOUSE STANDARD; PRT; 324 AA.

AC P01868;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE IG GAMMA-1 CHAIN C REGION.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80045036; PubMed=115593;
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
 RA Takahashi N., Mano Y.;
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
 gamma 1 chain gene.";
 RT Cell 18:559-568(1979).
 RL [2]
 RN [3]
 RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
 RX MEDLINE=80202559; PubMed=6769752;
 RA Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
 RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.;
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
 cloned in a bacterial plasmid.";
 RL Gene 9:87-97(1980).
 RN [3]
 RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
 RX MEDLINE=80012837; PubMed=113776;
 RA Rogers J., Clarke P., Salsler W.;
 RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin
 heavy chain.";
 RL Nucleic Acids Res. 6:3305-3321(1979).
 RN [4]
 RP SEQUENCE (MYELOMA PROTEIN MOPC 21).
 RX MEDLINE=78242288; PubMed=98524;
 RA Adetugbo K.;
 RT "Evolution of immunoglobulin subclasses. Primary structure of a
 murine myeloma gammal chain.";
 RL J. Biol. Chem. 253:6068-6075(1978).
 RN [5]
 RP DISULFIDE BONDS (MOPC 21).
 RX MEDLINE=73008889; PubMed=5073237;
 RA Svasti J., Milstein C.;
 RT "The disulphide bridges of a mouse immunoglobulin G1 protein.";
 RL Biochem J 126:837-850(1972).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; V00793; CAA24172.1; -
 CC EMBL; V00793; CAA24173.1; -
 CC EMBL; V00793; CAA24174.1; -
 CC EMBL; V00793; CAA24175.1; -
 CC EMBL; V00793; CAA24176.1; -

DR PIR; A02159; GIMS.
 DR GlycosuitedB; P01868;
 DR MGD; MGI:96446; Igh-4.
 DR InterPro; IPR003006; Iq_MHC.
 DR InterPro; IPR003597; Iq_cl.
 DR Pfam; PF00047; Iq; 3.
 DR SMART; SM00407; Iqcl; 2.
 DR PROSITE; PS00290; Iq_MHC; 1.
 KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
 KW Alternative splicing.
 FT NON_TER 1 97
 FT DOMAIN 1 97
 FT DOMAIN 98 110
 FT DOMAIN 111 217
 FT DOMAIN 218 324
 FT DISULFID 27 82
 FT DISULFID 102 102
 FT DISULFID 104 104
 FT DISULFID 107 107
 FT DISULFID 109 109
 FT DISULFID 138 198
 FT CARBOHYD 174 174
 FT FTID-CAR_000055.
 FT DISULFID 244 302
 FT MOD_RES 324 324
 FT CONFLICT 276 276
 FT CONFLICT 278 278
 FT SEQUENCE 324 AA; 35704 MW; A338812F3D1F2C93 CRC64;
 SQ

Query Match 33.1%; Score 48; DB 1; Length 324;
 Best Local Similarity 39.1%; Pred. No. 7;
 Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHFKPIVRSITK 24

DB 193 GKEFKCRVNSAAPPAPIEKTISK 215

RESULT 8

GCL_MOUSE STANDARD; PRT; 393 AA.

AC P01869;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE IG GAMMA-1 CHAIN C REGION, MEMBRANE-BOUND FORM.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80045036; PubMed=115593;
 RA Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
 RA Takahashi N., Mano Y.;
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
 gamma 1 chain gene.";
 RL Cell 18:559-568(1979).
 RN [2]
 RP SEQUENCE OF 323-393 FROM N.A.
 RX MEDLINE=82115295; PubMed=6804950;
 RA Tyler B.M., Cowman A.F., Gerondakis S.D., Adams J.M., Bernard O.;
 RT "mRNA for surface immunoglobulin gamma chains encodes a highly
 RT conserved transmembrane sequence and a 28-residue intracellular
 RT domain.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:2008-2012(1982).
 RN [3]
 RP SEQUENCE OF 323-366 FROM N.A.
 RX MEDLINE=82115295; PubMed=6799207;
 RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
 RA Eisenberg D., Wall R.;
 RT "Gene segments encoding transmembrane carboxyl termini of

```
RT Immunoglobulin gamma chains." ;
RL Cell 26:19-27(1981).
RN [4]
RP SEQUENCE OF 1-44 FROM N.A.
RX MEDLINE-82222190; PubMed-6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
  Immunoglobulin gamma chains." ;
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -1- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
  SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
  GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
  BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
  IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
  SEGMENT OF MU CHAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL; V00793; CAA24172.1; -
DR EMBL; V00793; CAA24173.1; -
DR EMBL; V00793; CAA24174.1; -
DR PIR; B02159; GIMSM.
DR MGD; MGI:96446; Igh-4.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IGH1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing; Transmembrane.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 37 82
FT DISULFID 102 102 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198
FT CARBOHYD 174 174 N-LINKED (GLCNAC. . .).
FT DISULFID 244 302
FT TRANSMEM 340 357
FT DOMAIN 358 393 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 393 AA; 43386 MW; 4CC88343B7A1CE27 CRC64;

Query Match 33.1%; Score 48; DB 1; Length 393;
Best Local Similarity 39.1%; Pred. No. 9.3;
Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
   |::||: ||||::|
DB 193 GKFKCRVNSAAPPAPIEKTISK 215

RESULT 9
GCA_RAT STANDARD; PRT; 322 AA.
AC P20760;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG GAMMA-2A CHAIN C REGION.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89232738; PubMed-3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family." ;
RL Gene 74:473-482(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M13804; AAA1376.1; ALT_INIT.
DR PIR; PS0019; PS0019.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00407; IGH1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 136 196
FT DISULFID 242 300
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 322 AA; 35186 MW; E8EA136A9DE01EDB CRC64;

Query Match 32.4%; Score 47; DB 1; Length 322;
Best Local Similarity 39.1%; Pred. No. 11;
Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPIVRSITK 24
   |::||: ||||::|
DB 191 GKTFCKVNSGAPAPIEKSISK 213

RESULT 10
DER2_HUMAN STANDARD; PRT; 345 AA.
AC Q9H3K2; O95894; Q9H0P2;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DERMAL PAPILLA DERIVED PROTEIN 2 (MYO21 PROTEIN).
GN DERP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Dermal papilla;
RA Ikeda A., Yamashita M., Yoshimoto M.;
RT "Molecular cloning of a dermal papilla derived gene." ;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Mei G., Yu W., Gibbs R.A.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Mao Y.M., Xie Y., Mu Z.M., Li Y., Huang Y.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
```

```

RN  SEQUENCE FROM N.A.
RP  TISSUE=Kidney;
RX  MEDLINE=21154917; PubMed=11230166;
RA  Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA  Ansgore W., Boecher M., Bloeker H., Bauersachs S., Blum H.,
RA  Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA  Lewes H.-W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA  Wambutt R., Korn B., Klein M., Poustka A.;
RT  Towards a catalog of human genes and proteins: sequencing and
RT  analysis of 500 novel complete protein coding human cDNAs.;
RL  Genome Res. 11:422-435(2001).
CC  -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC  -!- SIMILARITY: BELONGS TO THE BIL FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AB009685; BAA93049.1; -
DR  EMBL; AF131820; AAD20052.1; -
DR  EMBL; AF060923; AAG43135.1; -
DR  EMBL; AL136713; CAB66648.1; -
KW  Transmembrane.
FT  TRANSMEM 83 103 POTENTIAL.
FT  TRANSMEM 126 146 POTENTIAL.
FT  TRANSMEM 160 180 POTENTIAL.
FT  TRANSMEM 191 211 POTENTIAL.
FT  TRANSMEM 214 234 POTENTIAL.
FT  TRANSMEM 245 265 POTENTIAL.
FT  TRANSMEM 272 292 POTENTIAL.
FT  TRANSMEM 303 323 POTENTIAL.
FT  CONFLICT 4 4 A -> V (IN REF. 4).
FT  CONFLICT 64 64 E -> G (IN REF. 3).
FT  CONFLICT 74 74 I -> M (IN REF. 4).
FT  CONFLICT 118 118 Q -> R (IN REF. 4).
SQ  SEQUENCE 345 AA; 37205 MW; 808FAED86A9CD98E CRC64;

Query Match 32.1%; Score 46.5; DB 1; Length 345;
Best Local Similarity 51.9%; Pred. No. 14;
Matches 14; Conservative 2; Mismatches 8; Indels 3; Gaps 2;

QY 1 CGEGYQSRVDHPHPFK--PIVR-SITK 24
Db 8 CLRLPSRVFHPAFTKASPVVRNSITK 34
| ||| ||| | | | | | | | | | |
| ||| ||| ||| ||| ||| ||| |||

RESULT 11
SOR_METJA
ID SOR_METJA STANDARD; PRT; 116 AA.
AC Q58151;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PUTATIVE SUPEROXIDE REDUCTASE (EC 1.15.-.-) (SOR).
GN MJ0741.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,

Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
*Complete genome sequence of the methanogenic archaean, Methanococcus
jannaschii.;
Science 273:1058-1073(1996).
-!- FUNCTION: USES ELECTRONS FROM REDUCED NADP, BY WAY OF RUBREDOXIN
AND AN OXIDOREDUCTASE, TO CATALYZE THE REDUCTION OF SUPEROXIDE TO
HYDROGEN PEROXIDE (BY SIMILARITY).
-!- CATALYTIC ACTIVITY: O(2)- + 2 H(+) = H(2)O(2).
-!- COFACTOR: IRON (BY SIMILARITY).
-!- SIMILARITY: BELONGS TO THE DESULFOFERRODOXIN FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; U67520; AAB98735.1; -
DR  TIGR; MJ0741; -
DR  InterPro; IPR002742; Desulfoferrodox.
DR  Pfam; PF01880; Desulfoferrodox; 1.
DR  ProDom; PD006618; Desulfoferrodox; 1.
KW Hypothetical protein; Oxidoreductase; Electron transport; Iron;
KW Complete proteome.
FT METAL 20 20 IRON (BY SIMILARITY).
FT METAL 46 46 IRON (BY SIMILARITY).
FT METAL 52 52 IRON (BY SIMILARITY).
FT METAL 101 101 IRON (BY SIMILARITY).
FT METAL 104 104 IRON (BY SIMILARITY).
SQ SEQUENCE 116 AA; 13950 MW; 4CF2C76237DE0673 CRC64;

Query Match 31.7%; Score 46; DB 1; Length 116;
Best Local Similarity 43.5%; Pred. No. 5.3;
Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPFKPIVRSITK 24
Db 62 GDLYLARVDFTQFMKPEVKLMVK 84
| : ||| ||| ||| ||| : |
| : ||| ||| ||| ||| : |

RESULT 12
VNST_CVBQ
ID VNST_CVBQ STANDARD; PRT; 277 AA.
AC P18517;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-NOV-1990 (Rel. 16, Last annotation update)
DE 32 KDA NONSTRUCTURAL PROTEIN.
OS Bovine coronavirus (strain Quebec).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11133;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89345182; PubMed=2762160;
RX Cox G.J., Parker M.D., Babiuk L.A.;
RT "The sequence of cDNA of bovine coronavirus 32K nonstructural gene.";
RL Nucleic Acids Res. 17:5847-5847(1989).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X15445; CAA33485.1; -
```


Db 23 CGGYHQLCHTPHIDSSVIDSDEK 46
Search completed: March 4, 2002, 13:08:42
Job time: 639 sec

Query Match 31.7%; Score 46; DB 1; Length 370;
Best Local Similarity 87.5%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 12 PHEPKPIV 19
Db 110 PHEPKPLV 117

RESULT 15
MTF2_MOUSE
ID MTF2_MOUSE STANDARD; PRT; 373 AA.
AC Q02395;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE METAL-RESPONSE ELEMENT-BINDING TRANSCRIPTION FACTOR 2 (ZINC-REGULATED
DE FACTOR 1) (ZIRF1) (METAL-RESPONSE ELEMENT DNA-BINDING PROTEIN M96)
DE (FRAGMENT).
GN MTF2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Lymphoma;
RX MEDLINE=95290093; PubMed=7772254;
RA Inouye C., Remondelli P., Karin M., Elledge S.;
RT "Isolation of a cDNA encoding a metal response element binding protein
using a novel expression cloning procedure: the one hybrid system.";
RL DNA Cell Biol. 13:731-742(1994).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=97234834; PubMed=9173905;
RA Remondelli P., Leone A.;
RT "Interactions of the zinc-regulated factor (Zirf1) with the mouse
metallothionein 1a promoter.";
RL Biochem. J. 323:79-85(1997).
CC -!- FUNCTION: BINDS TO THE METAL-REGULATING-ELEMENT (MRE) OF
METALLOTHIONEIN 1A GENE PROMOTER. BINDING IS ZINC-DEPENDENT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: CONTAINS 2 PHD-TYPE ZINC FINGERS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S78454; AAC34714.1; -;
DR MGD; MGI:105050; Mtf2.
DR InterPro; IPR001965; PHD.
DR Pfam; PF00628; PHD; 2.
DR SMART; SM00249; PHD; 2.
KW DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;
KW Metal-binding.
FT NON_TER 1 1
FT ZN_FING 2 55 PHD-TYPE 1.
FT ZN_FING 101 153 PHD-TYPE 2.
FT DOMAIN 235 238 POLY-LYS.
SQ SEQUENCE 373 AA; 42003 MW; 50D2ECD13A5DE817 CRC64;

Query Match 31.7%; Score 46; DB 1; Length 373;
Best Local Similarity 33.3%; Pred. No. 17;
Matches 8; Conservative 3; Mismatches 13; Indels 0; Gaps 0;
QY 1 CGEGYOSRVDHPHPKPIVRSITK 24
||:|| ||::| |

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:08:11 ; Search time 79.15 Seconds
(without alignments)
46.201 Million cell updates/sec

Title: US-09-701-623C-7
Perfect score: 145
Sequence: 1 CGEGYQSRVDHPHPRPVRSYK 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_17.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query % Match | Length | ID | Description |
|------------|-------|---------------|--------|-----------|---------------------|
| 1 | 62 | 42.8 | 2421 | 5 Q9VM88 | Q9vm88 drosophila |
| 2 | 54 | 37.2 | 648 | 5 Q9NKKD7 | Q9nkd7 drosophila |
| 3 | 54 | 37.2 | 701 | 5 Q9VJU4 | Q9vju4 drosophila |
| 4 | 52 | 35.9 | 217 | 7 Q30827 | Q30827 ovis aries |
| 5 | 51 | 35.2 | 245 | 7 Q31270 | Q31270 rattus norv |
| 6 | 51 | 35.2 | 261 | 7 Q9TQA7 | Q9tqa7 rattus norv |
| 7 | 51 | 35.2 | 264 | 7 Q9TQA5 | Q9tqa5 rattus norv |
| 8 | 50 | 34.5 | 140 | 12 Q99A59 | Q99a59 bovine vira |
| 9 | 50 | 34.5 | 142 | 7 Q95536 | Q95536 pongo pygma |
| 10 | 50 | 34.5 | 314 | 11 Q9ESH2 | Q9esh2 rattus norv |
| 11 | 49 | 33.8 | 125 | 11 Q9CZ59 | Q9cz59 mus musculus |
| 12 | 49 | 33.8 | 173 | 7 Q95594 | Q95594 gallus gall |
| 13 | 49 | 33.8 | 273 | 7 Q98194 | Q98194 acrocephalu |
| 14 | 49 | 33.8 | 338 | 7 Q31412 | Q31412 gallus gall |
| 15 | 49 | 33.8 | 344 | 7 Q46792 | Q46792 gallus gall |
| 16 | 49 | 33.8 | 351 | 7 Q95593 | Q95593 gallus gall |
| 17 | 49 | 33.8 | 355 | 7 Q46788 | Q46788 gallus gall |
| 18 | 49 | 33.8 | 355 | 7 Q46789 | Q46789 gallus gall |
| 19 | 49 | 33.8 | 355 | 7 Q46790 | Q46790 gallus gall |

| | | | | | | |
|----|------|------|------|----|--------|--------------------|
| 20 | 49 | 33.8 | 355 | 7 | Q46791 | Q46791 gallus gall |
| 21 | 49 | 33.8 | 355 | 7 | Q31400 | Q31400 gallus gall |
| 22 | 49 | 33.8 | 355 | 7 | Q95601 | Q95601 gallus gall |
| 23 | 49 | 33.8 | 355 | 7 | Q9GIP6 | Q9gip6 gallus gall |
| 24 | 49 | 33.8 | 355 | 13 | Q73901 | Q73901 gallus gall |
| 25 | 49 | 33.8 | 355 | 13 | Q73904 | Q73904 gallus gall |
| 26 | 49 | 33.8 | 359 | 7 | Q98192 | Q98192 acrocephalu |
| 27 | 49 | 33.8 | 359 | 7 | Q98193 | Q98193 acrocephalu |
| 28 | 49 | 33.8 | 360 | 7 | Q98187 | Q98187 acrocephalu |
| 29 | 49 | 33.8 | 360 | 7 | Q98188 | Q98188 acrocephalu |
| 30 | 49 | 33.8 | 360 | 7 | Q98191 | Q98191 acrocephalu |
| 31 | 49 | 33.8 | 1208 | 5 | Q21117 | Q21117 caenorhabdi |
| 32 | 48.5 | 33.4 | 550 | 5 | Q17145 | Q17145 lucilia cup |
| 33 | 48.5 | 33.4 | 597 | 5 | Q76664 | Q76664 caenorhabdi |
| 34 | 48 | 33.1 | 93 | 7 | Q31260 | Q31260 rattus norv |
| 35 | 48 | 33.1 | 196 | 7 | P79590 | P79590 rattus norv |
| 36 | 48 | 33.1 | 237 | 6 | Q77499 | Q77499 sus scrofa |
| 37 | 48 | 33.1 | 266 | 7 | Q31071 | Q31071 sus scrofa |
| 38 | 48 | 33.1 | 266 | 7 | Q31072 | Q31072 sus scrofa |
| 39 | 48 | 33.1 | 437 | 11 | Q9R1A4 | Q9r1a4 mus musculu |
| 40 | 48 | 33.1 | 463 | 11 | Q99LC4 | Q99lc4 mus musculu |
| 41 | 48 | 33.1 | 1735 | 4 | Q9HBL0 | Q9hbl0 homo sapien |
| 42 | 47.5 | 32.8 | 1198 | 3 | Q9P5J0 | Q9p5j0 neurospora |
| 43 | 47 | 32.4 | 533 | 3 | Q9R3Y0 | Q9r3y0 streptomyce |
| 44 | 47 | 32.4 | 1413 | 5 | Q9XWR0 | Q9xwr0 caenorhabdi |
| 45 | 47 | 32.4 | 2254 | 10 | Q9LND2 | Q9lnd2 arabidopsis |

ALIGNMENTS

RESULT 1

Q9VM88 ID Q9VM88 PRELIMINARY: PRT: 2421 AA.
AC Q9VM88;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE CG11321 PROTEIN.
GN CG11321.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Adayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadien E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry J., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL SMCB 287:2185-2195(2000).
DR EMBL; AE003615; AAF52435.1; -
DR FlyBase; FBgn0031857; CG11321.
DR InterPro; IPR002106; AA_rRNA_ligase_II.
DR InterPro; IPR001870; Znf-RanBP.
DR SMART; SM001841; Znf-Ring.
DR SMART; SM00184; RING; 2.
DR SMART; SM00547; Znf-RBZ; 1.
DR PROSITE; PS00339; AA_rRNA_LIGASE_II.2; UNKNOWN_1.
SQ SEQUENCE 2421 AA; 269655 MW; 6F15EPD9128F13E0 CRC64;

Query Match 42.8%; Score 62; DB 5; Length 2421;
Best Local Similarity 52.9%; Pred. No. 1.6;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVHPHPKPI 18
| | | | | | | | | |
Db 211 GPGYMSQHHPPYPPV 227

RESULT 2
Q9NKD7 PRELIMINARY; PRT; 648 AA.
ID Q9NKD7
AC Q9NKD7
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHEICAL 69.4 KDA PROTEIN.
GN BG:DS00180.14 OR CG18146.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, CN BW SP;
RX MEDLINE=9940:001; PubMed=10471707;
RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA Celniker S., Rubin G.M.;
RA "An exploration of the sequence of a 2.9-Mb region of the genome of
RT *Drosophila melanogaster*: the Adh region.";
RL Genetics 153:179-219(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Y, CN BW SP;
RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,
RA Butenhoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,
RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lomont M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sethi H., Snair E., Swirskas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF003408; AAF44845.1; -

DR FlyBase; FBgn0028939; BG:DS00180.14.
DR InterPro; IPR000561; EGF-like.
DR SMART; SM00181; EGF; 10.
DR SMART; SM00001; EGF-like; 2.
DR PROSITE; PS01186; EGF_2; UNKNOWN_6.
KW Hypothetical protein.
SQ SEQUENCE 648 AA; 69419 MW; 3F16E0EBF4B94CF8 CRC64;

Query Match 37.2%; Score 54; DB 5; Length 648;
Best Local Similarity 69.2%; Pred. No. 6.9;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGGYQSRVHPDH 13
| | | | | | | | | |
Db 45 CGGYVSRKDHGH 57

RESULT 3
Q9VJU4 PRELIMINARY; PRT; 701 AA.
ID Q9VJU4
AC Q9VJU4
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CG18146 PROTEIN.
GN BG:DS00180.14 OR CG18146.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Ye H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";

```

RL Science 287:2185-2195(2000).
DR EMBL; AE003642; AAF53366.1; -.
DR HSSP; P02876; SWGA.
DR FlyBase; FBgn0028939; BG:DS00180.14.
DR InterPro; IPR000561; EGF-like.
DR SMART; SM00181; EGF; 10.
DR SMART; SM00001; EGF-like; 2.
DR PROSITE; PS01186; EGF_2; UNKNOWN_6.
SQ SEQUENCE 701 AA; 75366 MW; 95159C04C9C09BD4 CRC64;

Query Match 37.2%; Score 54; DB 5; Length 701;
Best Local Similarity 69.2%; Pred. No. 7.4;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPH 13
Db 98 CGKGYVSRKDHG 110
||| ||| |||

RESULT 4
Q30827 PRELIMINARY; PRT; 217 AA.
ID Q30827
AC Q30827
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE MHC CLASS LL DOB GENE (LEADER SEQUENCE) (FRAGMENT).
GN OVAR-DOB.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FINNISH LANDRACE; TISSUE-BLOOD;
RX MEDLINE=96128257; PubMed=8537127;
RA Wright H., Redmond J., Ballingali K.T.;
RT "The sheep orthologue of the HLA-DOB gene.";
RL Immunogenetics 43:76-79(1996).
CC -1- MISCELLANEOUS: FOUND N-TERMINAL TO IMMUNOGLOBULIN AND MAJOR
CC HISTOCOMPATIBILITY COMPLEX DOMAIN (BY SIMILARITY).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
CC -1- SIMILARITY: TO CLASS II HISTOCOMPATIBILITY ANTIGEN, BETA CHAIN,
CC BETA-1 DOMAIN INTERPRO FAMILY.
CC EMBL; M12382; AAA41610.1; -.
DR HSSP; P06343; IIAK.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR000353; MHC_II_beta.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF00969; MHC_II_beta; 1.
DR ProDom; PD000328; MHC_II_beta; 1.
DR SMART; SM00407; IGcl; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Glycoprotein; MHC; MHC II; Transmembrane.
FT NON_TER 1
SQ SEQUENCE 245 AA; 28069 MW; 6B21DADB2A4A299D CRC64;

Query Match 35.2%; Score 51; DB 7; Length 245;
Best Local Similarity 47.1%; Pred. No. 7.4;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHFKPI 18
Db 176 GEVYTCQVEHPSPV 192
||| :||| |||

RESULT 6
Q9TOA7 PRELIMINARY; PRT; 261 AA.
ID Q9TOA7
AC Q9TOA7
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE CLASS II MHC RT1.D(A) BETA CHAIN PRECURSOR.
DE RT1.D.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ACI;
RX MEDLINE=99299366; PubMed=10369938;
RA Tian L., Wang M., Yu J., Kahan B.D., Stepkowski S.M.;
RT "Nucleotide sequences of three distinct complementary DNA clones
RT encoding rat class II major histocompatibility complex RT1.D beta-
RT chain proteins.";
RL Immunogenetics 49:735-737(1999).
CC -1- MISCELLANEOUS: FOUND N-TERMINAL TO IMMUNOGLOBULIN AND MAJOR
CC HISTOCOMPATIBILITY COMPLEX DOMAIN (BY SIMILARITY).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX

```

```

CC          DOMAIN.
CC  -1- SIMILARITY: TO CLASS II HISTOCOMPATIBILITY ANTIGEN, BETA CHAIN,
CC      BETA-1 DOMAIN INTERPRO FAMILY.
CC      EMBL: AF084932; AAD39082.1; -.
CC      InterPro: IPR003597; Ig_cl.
CC      InterPro: IPR003006; Ig_MHC.
CC      Pfam: PF00047; Ig; 1.
CC      ProDom: PD000328; MHC_II_beta; 1.
CC      SMART: SM00407; IG1; 1.
CC      PROSITE: PS00290; IG_MHC; UNKNOWN_1.
CC      Glycoprotein: MHC; MHC II; Signal; Transmembrane.
CC      SIGNAL 1 26 POTENTIAL.
CC      FT CHAIN 27 261 CLASS II MHC RT1.D(A) BETA CHAIN.
CC      SQ SEQUENCE 261 AA; 29496 MW; 48E220A69B374B3D CRC64;

Query Match 35.2%; Score 51; DB 7; Length 261;
Best Local Similarity 47.1%; Pred. No. 7.9;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPI 18
Db 192 GEVYTCQVEHPSPV 208

RESULT 7
Q9TQA5 PRELIMINARY; PRT; 264 AA.
AC Q9TQA5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CLASS II MHC RT1.D(U) BETA CHAIN PRECURSOR.
GN RT1.D.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR FURTH;
RX MEDLINE=99299366; PubMed=10369938;
RA Tian L., Wang M., Yu J., Kahan B.D., Stepkowski S.M.;
RT "Nucleotide sequences of three distinct complementary DNA clones
RT encoding rat Class II major histocompatibility complex RT1.D beta-
RT chain proteins."
RL Immunogenetics 49:735-737(1999).
CC -1- MISCELLANEOUS: FOUND N-TERMINAL TO IMMUNOGLOBULIN AND MAJOR
CC HISTOCOMPATIBILITY COMPLEX DOMAIN (BY SIMILARITY).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
CC -1- SIMILARITY: TO CLASS II HISTOCOMPATIBILITY ANTIGEN, BETA CHAIN,
CC      BETA-1 DOMAIN INTERPRO FAMILY.
CC      EMBL: AF084934; AAD39084.1; -.
CC      InterPro: IPR003597; Ig_cl.
CC      InterPro: IPR003006; Ig_MHC.
CC      InterPro: IPR000353; MHC_II_beta.
CC      Pfam: PF00047; Ig; 1.
CC      ProDom: PD000328; MHC_II_beta; 1.
CC      SMART: SM00407; IG1; 1.
CC      PROSITE: PS00290; IG_MHC; UNKNOWN_1.
CC      Glycoprotein: MHC; MHC II; Signal; Transmembrane.
CC      SIGNAL 1 26 POTENTIAL.
CC      FT CHAIN 27 264 CLASS II MHC RT1.D(U) BETA CHAIN.
CC      SQ SEQUENCE 264 AA; 30138 MW; 3C888533514F531F CRC64;

Query Match 35.2%; Score 51; DB 7; Length 264;
Best Local Similarity 47.1%; Pred. No. 7.9;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

```

```

QY 2 GEGYQSRVDHPHPKPI 18
Db 195 GEVYTCQVEHPSPV 211

RESULT 8
Q99A59 PRELIMINARY; PRT; 140 AA.
AC Q99A59;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE E2 GLYCOPROTEIN (FRAGMENT).
GN E2.
OS bovine viral diarrhoea virus strain 4998/89.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Pestivirus.
OX NCBI_TaxID=145184;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=4998/89;
RA Greiser-Wilke I.M.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=4998/89;
RA Tajima M.;
RT "Prevalence of genotypes 1 and 2 of bovine viral diarrhoea virus in
RT Lower Saxony, Germany."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ302959; CAC24757.1; -.
FT NON_TER 1 1
FT NON_TER 140 140
SQ SEQUENCE 140 AA; 15756 MW; 1A28D06B74391F74 CRC64;

Query Match 34.5%; Score 50; DB 12; Length 140;
Best Local Similarity 36.0%; Pred. No. 6;
Matches 9; Conservative 4; Mismatches 6; Indels 6; Gaps 1;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
Db 69 CGHFRKKEDLPHYP-----IGKC 87

RESULT 9
Q95536 PRELIMINARY; PRT; 142 AA.
AC Q95536;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE POPY-DRB5*0603 PROTEIN (FRAGMENT).
GN POPY-DRB5*0603.
OS Pongo pygmaeus pygmaeus (Bornean orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9602;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93177102; PubMed=8439728;
RA Schonbach C., Vincek V., Mayer W.E., Golubic M., O'Huigin C.,
RA Klein J.;
RT "Multiplication of Mhc-DRB5 loci in the orangutan: Implications for
RT the evolution of DRB haplotypes."
RL Mamm. Genome 4:159-170(1993).
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
CC EMBL: S56369; AAD13881.1; -.
CC HSSP: P06343; IIAK.
CC InterPro: IPR003597; Ig_cl.
CC InterPro: IPR003006; Ig_MHC.

```

DR Pfam: PF00047; ig: 1.
 DR SMART: SM00407; IGcl: 1.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 142 AA; 15525 MW; 02C3C3AED44D063D CRC64;

Query Match 34.5%; Score 50; DB 7; Length 142;
 Best Local Similarity 47.1%; Pred. No. 6.1;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPI 18
 |||||:|:|:|:|:
 Db 73 GEGYTQVEHPSTVPL 89

RESULT 10

Q9ESH2 PRELIMINARY; PRT; 314 AA.
 AC Q9ESH2;
 DT 01-MAR-2001 (TEMBLrel. 16, Created)
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
 DE PRECONDITIONING-INDUCIBLE GENE 1 PROTEIN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 ON NCBI_TaxID=101116;
 RP SEQUENCE FROM N.A.
 RA Laser M., Li Y., Xu L., Darden A., Wu B.X., Hazard E.S. III,
 RA Crosson C., Ma J.X.;
 RT "Identification and characterization of a novel gene induced by
 RT ischemic preconditioning in the retina."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF175224; AAC09182.1; -
 SQ SEQUENCE 314 AA; 35838 MW; 4C64B70F7E909BDF CRC64;

Query Match 34.5%; Score 50; DB 11; Length 314;
 Best Local Similarity 43.8%; Pred. No. 13;
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 10 DHPHPKPIVRSITKC 25
 :||| | | : | : |
 Db 249 NHPHPPLMCSVHPK 264

RESULT 11

Q9CZ59 PRELIMINARY; PRT; 125 AA.
 AC Q9CZ59;
 DT 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
 DE 2810404006RIK PROTEIN.
 GN 2810404006RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RP SEQUENCE FROM N.A.
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohitsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL: AK012987; BAB28580.1; -
 DR MGD; MGI:1917376; 2810404006RIK.
 SQ SEQUENCE 125 AA; 14130 MW; 97B98DB40062AC9D CRC64;

Query Match 33.8%; Score 49; DB 11; Length 125;
 Best Local Similarity 53.3%; Pred. No. 7.5;
 Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 11 HPHFPKPIVRSITKC 25
 ||||| | : |
 Db 83 HPHFPILTPRLSC 97

RESULT 12

Q95594 PRELIMINARY; PRT; 173 AA.
 AC Q95594;
 DT 01-FEB-1997 (TEMBLrel. 02, Created)
 DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
 DE MINOR CLASS I GLYCOPROTEIN (FRAGMENT).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 ON NCBI_TaxID=9031;
 RP SEQUENCE FROM N.A.
 RC STRAIN-CC; TISSUE-SPLEEN;
 RA Wallny H., Avila D., Riegert P., Salomonsen J., Vilbois F., Wiles M.,
 RA Kaufman J.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
 CC IMMUNE SYSTEM (BY SIMILARITY).
 CC -!- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN) (BY SIMILARITY).
 CC -!- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
 CC DOMAIN.
 DR EMBL: Z54363; CAA91186.1; -
 DR InterPro: IPR003597; Ig_cl.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR001039; MHC_I.
 DR Pfam: PF00047; Ig; 1.
 DR Pfam: PF00129; MHC_I; 1.
 DR ProDom: PD000050; MHC_I; 1.
 DR SMART: SM00407; IGcl; 1.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
 KW Glycoprotein; MHC; Transmembrane.
 FT NON_TER 1
 FT NON_TER 173
 SQ SEQUENCE 173 AA; 19555 MW; E8EFBB56E410615B CRC64;

Query Match 33.8%; Score 49; DB 7; Length 173;
 Best Local Similarity 50.0%; Pred. No. 10;
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPHPKPI 17
 | : | | | | : | : |
 Db 149 GDKYQCRVEHASUPQ 164

```

different rates. Chicken B-F and beta 2-microglobulin sequences reveal
invariant surface residues."
J. Immunol. 148:1532-1546(1992).
-|- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
IMMUNE SYSTEM (BY SIMILARITY).
-|- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
MICROGLOBULIN) (BY SIMILARITY).
-|- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
DOMAIN.
EMBL; M84766; AAA67507.1; -.
HSP; P30685; IAE.
InterPro; IPR003597; Ig_cl.
InterPro; IPR003006; Ig_MHC.
InterPro; IPR001039; MHC_I.
Pfam; PF00047; Ig; 1.
Pfam; PF00129; MHC_I; 1.
ProDom; PD000050; MHC_I; 1.
SMART; SM00407; IGCI; 1.
PROSITE; PS00290; IG_MHC; UNKNOWN_1.
Glycoprotein; MHC; Signal; Transmembrane.
NON_TER 1
FT SIGNAL <1 4 POTENTIAL.
FT CHAIN 5 338 MHC CLASS I-ALPHA.
SQ SEQUENCE 338 AA; 37573 MW; 2EAB3786059408F5 CRC64;

Query Match 33.8%; Score 49; DB 7; Length 338;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYSRVDPHPFKP 17
I: ||||| I:|
D 254 GDKYQCRVHEHSLPQP 269

RESULT 15
O46792 PRELIMINARY; PRT; 344 AA.
AC O46792;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MHC CLASS I GLYCOPROTEIN.
GN B-FIV.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B18 CONGENIC;
RX MEDLINE=98221119; PubMed=9553152;
RA Hunt H.D., Fulton J.E.;
RT "Analysis of polymorphisms in the major expressed class I locus (B-
FIV) of the chicken."
RL Immunogenetics 47:456-467(1998).
CC -|- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
IMMUNE SYSTEM (BY SIMILARITY).
CC -|- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
MICROGLOBULIN) (BY SIMILARITY).
CC -|- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
DOMAIN.
EMBL; AF013496; AAC17593.1; -.
HSP; P30685; IAE.
InterPro; IPR003597; Ig_cl.
InterPro; IPR003006; Ig_MHC.
InterPro; IPR001039; MHC_I.
Pfam; PF00047; Ig; 1.
Pfam; PF00129; MHC_I; 1.
ProDom; PD000050; MHC_I; 1.
SMART; SM00407; IGCI; 1.
SQ SEQUENCE 338 AA; 37573 MW; 2EAB3786059408F5 CRC64;

Query Match 33.8%; Score 49; DB 7; Length 273;
Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 EGYQSRVDHPHPFKP 17
I: ||||| I:|
D 191 EGYQSRVHPGMPKP 205

RESULT 14
Q31412 PRELIMINARY; PRT; 338 AA.
AC Q31412;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MHC CLASS I-ALPHA PRECURSOR (FRAGMENT).
GN B-F.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H.B19; TISSUE=BONE MARROW;
RX MEDLINE=92166395; PubMed=1538136;
RA Kaufman J., Andersen R., Avila D., Engberg J., Lambris J.,
RA Salomonsen J., Welinder K., Skjold K.;
RT "Different features of the MHC class I heterodimer have evolved at

```

KW Glycoprotein; MHC; Transmembrane.
SQ SEQUENCE 344 AA; 37923 MW; 48A50E061A8F965C CRC64;

Query Match 33.8%; Score 49; DB 7; Length 344;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 GEGYQSRVDHPFPKP 17
|: ||||| |:
Db 271 GDKYQCRVEHASLPQP 286

Search completed: March 4, 2002, 13:08:12
Job time: 694 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:56:35 ; Search time 38.04 Seconds
(without alignments)
14.789 Million cell updates/sec

Title: US-09-701-623c-7
Perfect score: 145
Sequence: 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 103 | 71.0 | 25 | 3 | US-09-100-414B-95 |
| 2 | 103 | 71.0 | 25 | 4 | US-09-303-323-95 |
| 3 | 103 | 71.0 | 42 | 3 | US-09-100-414B-98 |
| 4 | 103 | 71.0 | 42 | 3 | US-09-100-414B-99 |
| 5 | 103 | 71.0 | 42 | 3 | US-09-100-414B-100 |
| 6 | 103 | 71.0 | 42 | 4 | US-09-303-323-98 |
| 7 | 103 | 71.0 | 42 | 4 | US-09-303-323-99 |
| 8 | 103 | 71.0 | 42 | 4 | US-09-303-323-100 |
| 9 | 103 | 71.0 | 45 | 3 | US-09-100-414B-101 |
| 10 | 103 | 71.0 | 45 | 4 | US-09-303-323-101 |
| 11 | 103 | 71.0 | 46 | 3 | US-09-100-414B-96 |
| 12 | 103 | 71.0 | 46 | 3 | US-09-100-414B-97 |
| 13 | 103 | 71.0 | 46 | 4 | US-09-303-323-96 |
| 14 | 103 | 71.0 | 46 | 4 | US-09-303-323-97 |
| 15 | 103 | 71.0 | 63 | 3 | US-09-100-414B-102 |
| 16 | 103 | 71.0 | 63 | 4 | US-09-303-323-102 |
| 17 | 85 | 58.6 | 561 | 3 | US-09-192-545-2 |
| 18 | 82 | 56.6 | 426 | 1 | US-08-336-583-2 |
| 19 | 82 | 56.6 | 426 | 5 | PCT-US95-13795-2 |
| 20 | 80 | 55.2 | 106 | 2 | US-08-232-539D-54 |
| 21 | 80 | 55.2 | 113 | 2 | US-08-232-539D-56 |
| 22 | 69.5 | 47.9 | 119 | 2 | US-08-464-025A-1 |
| 23 | 62 | 42.8 | 22 | 2 | US-08-232-539D-19 |
| 24 | 62 | 42.8 | 24 | 2 | US-08-232-539D-20 |
| 25 | 62 | 42.8 | 56 | 2 | US-08-232-539D-18 |
| 26 | 53.5 | 36.9 | 118 | 3 | US-08-466-151-1 |
| 27 | 53 | 36.6 | 331 | 2 | US-08-646-981-17 |

US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-95

ALIGNMENTS

RESULT 1

US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-95

Query Match 71.0%; Score 103; DB 3; Length 25;
Best Local Similarity 68.0%; Pred. No. 1.5e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Oy 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Db 1 CGETYQSRVTHPLPALMRSTTKC 25

RESULT 2
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 71.0%; Score 103; DB 4; Length 25;
Best Local Similarity 68.0%; Pred. No. 1.9e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 1 CGETYQSRVTHPHLPALMRSTRKC 25

RESULT 3
US-09-100-414B-98
; Sequence 98, Application US/09100414B
; Patent No. 6023468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows

; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-98

Query Match 71.0%; Score 103; DB 3; Length 42;
Best Local Similarity 68.0%; Pred. No. 2.7e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 18 CGETYQSRVTHPHLPALMRSTRKC 42

RESULT 4
US-09-100-414B-99
; Sequence 99, Application US/09100414B
; Patent No. 6023468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-99

Query Match 71.0%; Score 103; DB 3; Length 42;
Best Local Similarity 68.0%; Pred. No. 2.7e-09;

Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 5

US-09-100-414B-100
; Sequence 100, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-751-6849
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-100

Query Match 71.0%; Score 103; DB 3; Length 42;
Best Local Similarity 68.0%; Pred. No. 2.7e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 6

US-09-303-323-98
; Sequence 98, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-751-6849
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-98

Query Match 71.0%; Score 103; DB 4; Length 42;
Best Local Similarity 68.0%; Pred. No. 2.7e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 7

US-09-303-323-99
; Sequence 99, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-751-6849
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-99

Query Match 71.0%; Score 103; DB 4; Length 42;
Best Local Similarity 68.0%; Pred. No. 2.7e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 18 CGETYQSRVTHPLPRALMRSTTKC 42

RESULT 8

US-09-303-323-100
; Sequence 100, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-100

Query Match 71.0%; Score 103; DB 4; Length 42;
Best Local Similarity 68.0%; Pred. No. 2.7e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 18 CGETYQSRVTHPLPRALMRSTTKC 42

RESULT 9

US-09-100-414B-101
; Sequence 101, Application US/09100414B

; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-101

Query Match 71.0%; Score 103; DB 3; Length 45;
Best Local Similarity 68.0%; Pred. No. 2.9e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 21 CGETYQSRVTHPLPRALMRSTTKC 45

RESULT 10

US-09-303-323-101
; Sequence 101, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: 09/100,414
;; FILING DATE: 20-JUNE-1998
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Maria H. Lin
;; REGISTRATION NUMBER: 29,323
;; REFERENCE/DOCKET NUMBER: 1151-4157
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-758-4800
;; TELEFAX: 212-751-6849
;; INFORMATION FOR SEQ ID NO: 101:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 45 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-303-323-101

Query Match 71.0%; Score 103; DB 4; Length 45;
Best Local Similarity 68.0%; Pred. No. 2.9e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
Db 21 CGETYQSRVTHPHLPALMRSTTKC 45

RESULT 11
US-09-100-414B-96
; Sequence 96, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-96

Query Match 71.0%; Score 103; DB 3; Length 46;
Best Local Similarity 68.0%; Pred. No. 3e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25

Db 22 CGETYQSRVTHPHLPALMRSTTKC 46
RESULT 12
US-09-100-414B-97
; Sequence 97, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-97

Query Match 71.0%; Score 103; DB 3; Length 46;
Best Local Similarity 68.0%; Pred. No. 3e-09;
Matches 17; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
Db 22 CGETYQSRVTHPHLPALMRSTTKC 46

RESULT 13
US-09-303-323-96
; Sequence 96, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-303-323-96

Query Match 71.0%; Score 103; DB 4; Length 46;
Best Local Similarity 68.0%; Pred. No. 3e-09; Indels 5; Gaps 0;
Matches 17; Conservative 3; Mismatches 5; Indels 5; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 22 CGETYQSRVTHPHLPRALMRSTTKC 46

RESULT 14
US-09-303-323-97
; Sequence 97, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 46 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-303-323-97

Query Match 71.0%; Score 103; DB 4; Length 46;
Best Local Similarity 68.0%; Pred. No. 3e-09; Indels 5; Gaps 0;
Matches 17; Conservative 3; Mismatches 5; Indels 5; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 22 CGETYQSRVTHPHLPRALMRSTTKC 46

RESULT 15
US-09-100-414B-102
; Sequence 102, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 63 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-102

Query Match 71.0%; Score 103; DB 3; Length 63;
Best Local Similarity 68.0%; Pred. No. 4.2e-09; Indels 5; Gaps 0;
Matches 17; Conservative 3; Mismatches 5; Indels 5; Gaps 0;

QY 1 CGEGYQSRVDHPHPKPIVRSITKC 25
||| ||||| ||| | : || |||
Db 39 CGETYQSRVTHPHLPRALMRSTTKC 63

Search completed: March 4, 2002, 12:56:35
Job time: 257 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:58:21 ; Search time 42.32 Seconds
(without alignments)
44.999 Million cell updates/sec

Title: US-09-701-623C-8
Perfect score: 140
Sequence: 1 CGYQSVIVDRPDPFKPIVRSITL 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|----------|---------------------|
| 1 | 116 | 82.9 | 423 | 1 EHMS | Ig epsilon chain C |
| 2 | 108 | 77.1 | 107 | 2 I68730 | IgE chain C3 regio |
| 3 | 108 | 77.1 | 107 | 2 I68726 | IgE chain C3 regio |
| 4 | 108 | 77.1 | 388 | 1 EHMS | Ig epsilon chain C |
| 5 | 108 | 77.1 | 548 | 2 S38864 | Ig epsilon chain C |
| 6 | 85 | 60.7 | 429 | 1 EHRT | Ig epsilon chain C |
| 7 | 49 | 35.0 | 165 | 2 T31043 | hypothetical prote |
| 8 | 49 | 35.0 | 433 | 2 C75354 | conserved hypotet |
| 9 | 49 | 35.0 | 705 | 2 A75026 | phosphoribosylform |
| 10 | 48.5 | 34.6 | 363 | 1 BVBK2 | MAK32 protein - ye |
| 11 | 48 | 34.3 | 169 | 1 S12142 | cytochrome-c oxida |
| 12 | 48 | 34.3 | 294 | 2 T32293 | phosphate butyrylt |
| 13 | 48 | 34.3 | 1259 | 2 T06521 | pitillysin (EC 3.4 |
| 14 | 47 | 33.6 | 169 | 1 A35209 | cytochrome-c oxida |
| 15 | 47 | 33.6 | 241 | 2 E71532 | probable porphobil |
| 16 | 46.5 | 33.2 | 210 | 2 C75399 | hypothetical prote |
| 17 | 46 | 32.9 | 351 | 2 B70801 | probable fadE36 pr |
| 18 | 46 | 32.9 | 426 | 2 I36948 | Ig epsilon-chain - |
| 19 | 46 | 32.9 | 705 | 2 A71211 | probable phosphori |
| 20 | 46 | 32.9 | 715 | 2 S4628 | hypothetical prote |
| 21 | 45 | 32.1 | 172 | 2 S75440 | adenine phosphorib |
| 22 | 45 | 32.1 | 294 | 1 S72706 | probable enoyl-CoA |
| 23 | 45 | 32.1 | 910 | 2 H85868 | NADH dehydrogenase |
| 24 | 45 | 32.1 | 921 | 2 A65000 | NADH dehydrogenase |
| 25 | 45 | 32.1 | 921 | 2 G02326 | transcription fact |
| 26 | 44.5 | 31.8 | 62 | 1 SBHUP | statherin precursor |
| 27 | 44.5 | 31.8 | 222 | 2 S38779 | beta-amylase (EC 3 |
| 28 | 44.5 | 31.8 | 335 | 2 T16904 | hypothetical prote |
| 29 | 44.5 | 31.8 | 1056 | 2 S55151 | probable membrane |

| | | | | | |
|----|------|------|------|----------|--------------------|
| 30 | 44 | 31.4 | 46 | 2 D49281 | pol protein - simi |
| 31 | 44 | 31.4 | 169 | 1 OL804 | cytochrome-c oxida |
| 32 | 44 | 31.4 | 282 | 2 T39767 | translation initia |
| 33 | 44 | 31.4 | 587 | 1 KSKVAO | L-ascorbate oxidas |
| 34 | 44 | 31.4 | 699 | 2 B72775 | probable DNA repli |
| 35 | 44 | 31.4 | 706 | 2 D83112 | elongation factor |
| 36 | 44 | 31.4 | 1058 | 2 T30580 | p-type ATPase - sl |
| 37 | 43.5 | 31.1 | 143 | 2 H72296 | sugar-phosphate is |
| 38 | 43.5 | 31.1 | 257 | 2 S74478 | hypothetical prote |
| 39 | 43.5 | 31.1 | 305 | 2 T28063 | hypothetical prote |
| 40 | 43.5 | 31.1 | 332 | 2 E72597 | probable alcohol d |
| 41 | 43 | 30.7 | 54 | 2 G61492 | ovomucoid (pSTI-ty |
| 42 | 43 | 30.7 | 54 | 2 B31443 | ovomucoid, third d |
| 43 | 43 | 30.7 | 54 | 2 H31436 | ovomucoid, third d |
| 44 | 43 | 30.7 | 54 | 2 A31439 | ovomucoid, third d |
| 45 | 43 | 30.7 | 54 | 2 B31436 | ovomucoid, third d |

ALIGNMENTS

RESULT 1
EHMS
Ig epsilon chain C region (version 2) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Aug-1996
C:Accession: A02145
R:ishida, N.; Ueda, S.; Hayashida, H.; Miyata, T.; Honjo, T.
EMBO J. 1, 1117-1123, 1982
A:Title: The nucleotide sequence of the mouse immunoglobulin epsilon gene: comparison
A:Reference number: A90966; MUID:84236092
A:Accession: A02145
A:Molecule type: DNA
A:Residues: 1-423 <ISH>
A:Note: the sequence was determined from the germline gene
C:Genetics: 91/1; 199/1; 307/1
A:Introns: 91/1; 199/1; 307/1
C:Complex: An immunoglobulin heterotrimer subunit consists of two identical light (chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotrimer; immunoglobulin
F:16-77/Domain: immunoglobulin homology <IMM1>
F:115-183/Domain: immunoglobulin homology <IMM2>
F:220-288/Domain: immunoglobulin homology <IMM3>
F:325-396/Domain: immunoglobulin homology <IMM4>
F:325-75,122-181,227-286,332-394/Disulfide bonds: #status predicted
F:43,84,167,239,262,417/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 82.9%; Score 116; DB 1; Length 423;
Best Local Similarity 91.3%; Pred. No. 5.4e-10;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYGQSVIVDRPDPFKPIVRSITL 24
|||||:|||||
DB 281 GYGQSVIVDRPDPFKPIVRSITL 303

RESULT 2
I68730
IgE chain C3 region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C:Accession: I68730
R:Shinkai, Y.; Nakauchi, H.; Honjo, T.; Okumura, K.
Immunogenetics 27, 288-292, 1988
A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic aci
A:Reference number: I54443; MUID:88152907
A:Accession: I68730
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-107 <RES>
A:Cross-references: GB:M22933; NID:g194464; PIDN:AAA37915.1; PID:g194469

C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:22-90/Domain: immunoglobulin homology <IMM>

Query Match 77.1%; Score 108; DB 2; Length 107;
Best Local Similarity 90.9%; Pred. No. 2e-09;
Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GYGQSIQVDRDPFKPIVRSIT 23
||||| ||| |||||||||

Db 83 GYGQCIQVDRDPFKPIVRSIT 104

RESULT 3

168726

IgE chain C3 region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000

C:Accession: 168726

R:Shinkai, Y.; Nakauchi, H.; Honjo, T.; Okumura, K.

Immunogenetics 27, 288-292, 1988

A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic acid s

A:Reference number: 154443; MUID:88152907

A:Accession: 168726

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-107 <RES>

A:Cross-references: GB:M22930; NID:g194455; PIDN:AAA37911.1; PID:g194460

C:Superfamily: immunoglobulin C region; immunoglobulin homology

F:22-90/Domain: immunoglobulin homology <IMM>

Query Match

Best Local Similarity 77.1%; Score 108; DB 2; Length 107;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GYGQSIQVDRDPFKPIVRSIT 23

||||| ||| |||||||||

Db 83 GYGQCIQVDRDPFKPIVRSIT 104

RESULT 4

EHMS

Ig epsilon chain C region (version 1) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 16-Jul-1999

C:Accession: A02144

R:Liu, F.T.; Albrandt, K.; Sutcliffe, J.G.; Katz, D.H.

Proc. Natl. Acad. Sci. U.S.A. 79, 7852-7856, 1982

A:Title: Cloning and nucleotide sequence of mouse immunoglobulin epsilon chain cDNA.

A:Reference number: A02144; MUID:8311774

A:Accession: A02144

A:Molecule type: mRNA

A:Residues: 1-388 <LIU>

A:Cross-references: GB:J00476; NID:g194875; PIDN:AAA38085.1; PID:g387220

C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap

hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la

C:Superfamily: immunoglobulin C region; immunoglobulin homology

C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin

F:1-44/Domain: immunoglobulin homology (fragment) <IM1>

F:81-149/Domain: immunoglobulin homology <IM2>

F:186-254/Domain: immunoglobulin homology <IM3>

F:290-361/Domain: immunoglobulin homology <IM4>

F:10,51,62,133,205,228,332,382/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match

Best Local Similarity 77.1%; Score 108; DB 1; Length 388;

Matches 20; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GYGQSIQVDRDPFKPIVRSIT 23

||||| ||| |||||||||

Db 247. GYGQCIQVDRDPFKPIVRSIT 268

A;Accession: S16114
A;Molecule type: mRNA
A;Residues: 1-169 <ARC>
C;Genetics:
A;Gene: COXIV
A;Genome: nuclear

C:Keywords: membrane-associated complex; mitochondrion; oxidative phosphorylation; oxidoreductase; F₁F₀-ATP synthase
F:1-22/Domain: transit peptide (mitochondrion) #status predicted <NAP>
F:3-169/Product: cytochrome-c oxidase chain IV #status predicted <NAP>

Query Match 34.3%; Score 48; DB 1; Length 169;
Best Local Similarity 36.4%; Pred. No. 5.6;
Matches 8; Conservative 5; Mismatches 9; Indels

Qy 3 YGYQSI VDRP DFPKPI VRSITL 24
| : : | | | : | :
Db 33 YAFPTYAD RRDYPLPDVAHV TM 54

RESULT 12
G72293 phosphate butyryltransferase - Thermotoga maritima (strain MSB9)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: G72293
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.

Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing of *Mycobacterium thermophilus*
A:Reference number: A72200; MUID:99287316
A:Accession: G72293
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-294
A:Cross-references: GB:AE001770; GB:AE000512; NID:g4981658; PIDN:AD36206.1; PID:g498167
A:Experimental source: strain MSB8
C:Genetics:
C:Gene: TM1130
C:Superfamily: phosphate acetyltransferase

Query Match 34.3%; Score 48; DB 2; Length 294;
Best Local Similarity 46.7%; Pred. No. 10;
Matches 7; Conservative 6; Mismatches 2; Indels

QY 7 SIVDRPDEPKPIVRS 21
|::|::|::|
Db 122 SVMETPDEPRPLIIS 136

RESULT 13
T06521
pitrilysin (EC 3.4.24.55) - garden pea
C.Species: Pisum sativum (garden pea)
C.Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C.Accession: T06521
R:Vanderveere, P.S.; Bennett, T.M.; Blong, J.E.; Lampka, G.K.
Proc. Natl. Acad. Sci. U.S.A. 92, 7177-7181, 1995
A:Title: A chloroplast processing enzyme involved in precursor maturation shares a zinc-
A:Reference number: Z15733; MUID:95365331
A:Accession: T06521

C;Function:
A;Description: catalyzes proteolytic removal of chloroplast transit peptides
C;Keywords: hydrolase; metalloproteinase; zinc

Query Match 34.3%; Score 48; DB 2; Length 1259;
Best Local Similarity 38.1%; Pred. NO. 52;
Matches 8; Conservative 5; Mismatches 8; Indels

| Qy | 5 | YQSI | VR | DP | DF | PK | PI | VR | SI | TL | 25 |
|----|----|------|----|------|----|----|-----|-----|----|----|----|
| | | | | : | : | | | : | : | | |
| Db | 79 | YSSV | LS | QPTA | VP | VR | QSC | TSC | | | 99 |

RESULT 14

A35209 cytochrome-c oxidase (EC 1.9.3.1) chain IV precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 11-Jun-1999
C:Accession: A35209; S12724; S04709; S04593; S14190; S65373
R:Yamada, M.; Amuro, N.; Goto, Y.; Okazaki, T.
J. Biol. Chem. 265, 7687-7692, 1990
A:Title: Structural organization of the rat cytochrome c oxidase subunit IV gene.
A:Reference number: A35209; MUID:90237079
A:Accession: A35209
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-169 <YAM>
A:Cross-References: GB:J05425; NID:g203516; PIDN:AAA40949.1; PID:g203517
R:Amuro, N.; Yamada, M.; Goto, Y.; Okazaki, T.
Nucleic Acids Res. 18, 3992, 1990
A:Title: Complete nucleotide sequence of the gene encoding rat cytochrome c oxidase
A:Reference number: S12724; MUID:90326528

C:Function: the cytochrome-c oxidase complex catalyzes the oxidation of four molecules from the mitochondrial matrix producing two molecules of water and lowering the concentration of protons in the matrix

Query Match 33.6%; Score 47; DB 1; Length 169;
Best Local Similarity 45.5%; Pred. No. 7.9;
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

| | | | | | | | | | | |
|----|----|--------|----|----|----|----|----|----|-----|----|
| Qy | 3 | YGYQSI | VD | RP | DF | PK | PI | VS | ITL | 24 |
| | | | | | | | | | | |
| Db | 33 | YALPSY | VD | RR | DY | PL | PD | VA | HKL | 54 |

RESULT 15
 E71532
 probable porphobilinogen deaminase - Chlamydia trachomatis (serotype D, strain UW3/Cx)
 C:Species: Chlamydia trachomatis
 C:Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
 C:Accession: E71532
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis serotype D, strain UW-3/Cx
 A:Reference number: A71570; MUID:99000809
 A:Accession: E71532
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-241 <ARN>
 A:Cross-references: GB:AE001273; PID:g332871
 A:Experimental source: serotype D, strain UW-3/Cx
 C:Genetics:
 A:Gene: hemC

Query Match 33.6%; Score 47; DB 2; Length 241;
Best Local Similarity 47.8%; Pred. No. 12;
Matches 11; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

| | | | | | | | | | | |
|----|----|----------|----|----|----|----|----|----|----|-----|
| Qy | 1 | CGYGYQSI | VD | RP | DF | PK | PI | VS | IT | 23 |
| | | | | | | | | | | |
| Db | 90 | CDLGIHSA | KD | LP | EN | PK | AT | VS | IT | 112 |

Search completed: March 4, 2002, 12:58:21
Job time: 248 sec

THIS PAGE BLANK (99716)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:08:42 ; Search time 24.65 Seconds
(without alignments)
37.185 Million cell updates/sec

Title: US-09-701-623c-8
Perfect score: 140
Sequence: 1 GYGYQSVDRDPFKPIVRISITLC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------|---------------------|
| 1 | 108 | 77.1 | 421 | 1 EPC_MOUSE | P06336 mus musculus |
| 2 | 85 | 60.7 | 429 | 1 EPC_RAT | P01855 rattus norv |
| 3 | 49 | 35.0 | 705 | 1 PURL_PYRAB | Q9UXW6 pyrococcus |
| 4 | 48.5 | 34.6 | 363 | 1 MK32_YEAST | P23060 saccharomyc |
| 5 | 48 | 34.3 | 144 | 1 COX4_PITPI | O46585 pithecia pi |
| 6 | 48 | 34.3 | 169 | 1 COX4_MOUSE | P19783 mus musculus |
| 7 | 48 | 34.3 | 294 | 1 PTA_THEMA | Q9X014 thermotoga |
| 8 | 47 | 33.6 | 169 | 1 COX4_RAT | P10888 rattus norv |
| 9 | 47 | 33.6 | 241 | 1 HEM3_CHLTR | O84301 chlamydia t |
| 10 | 46 | 32.9 | 658 | 1 KPCL_LYTP1 | O25378 lytechinus |
| 11 | 46 | 32.9 | 705 | 1 PURL_PYRHO | O59621 pyrococcus |
| 12 | 45 | 32.1 | 172 | 1 APT_SYNV3 | P73935 synchocyst |
| 13 | 45 | 32.1 | 294 | 1 ECHC_MYCLE | P53526 mycobacteri |
| 14 | 45 | 32.1 | 907 | 1 NUOG_ECOLI | P33602 escherichia |
| 15 | 45 | 32.1 | 907 | 1 NUOG_SALTY | P33900 salmonella |
| 16 | 45 | 32.1 | 925 | 1 NFE2_HUMAN | Q13469 homo sapien |
| 17 | 44.5 | 31.8 | 62 | 1 STAT_HUMAN | P02808 homo sapien |
| 18 | 44.5 | 31.8 | 222 | 1 AMYB_SECC | P30271 secale cere |
| 19 | 44.5 | 31.8 | 335 | 1 Y094_CAEFL | P41844 caenorhabdi |
| 20 | 44.5 | 31.8 | 1056 | 1 YNN2_YEAST | P53914 saccharomyc |
| 21 | 44 | 31.4 | 55 | 1 COX4_CEBAP | O46589 cebus apell |
| 22 | 44 | 31.4 | 137 | 1 COX4_PERPO | O46586 perodicticu |
| 23 | 44 | 31.4 | 144 | 1 COX4_AOTAZ | O46584 aotus azara |
| 24 | 44 | 31.4 | 169 | 1 COX4_BOVIN | P00423 bos taurus |
| 25 | 44 | 31.4 | 169 | 1 COX4_RABIT | Q9T118 oryctolagus |
| 26 | 44 | 31.4 | 282 | 1 IF34_SCHPO | P78795 schizosacch |
| 27 | 44 | 31.4 | 587 | 1 ASO_CUCSA | P14133 cucumis sat |
| 28 | 44 | 31.4 | 886 | 1 DSCI_MOUSE | P55849 mus musculus |
| 29 | 44 | 31.4 | 1058 | 1 PMAI_DICDI | P54679 dictyosteli |
| 30 | 43 | 30.7 | 54 | 1 IOVO_CITRAE | P05579 circus aeru |
| 31 | 43 | 30.7 | 54 | 1 IOVO_GYPCO | P05578 gyps coprot |
| 32 | 43 | 30.7 | 54 | 1 IOVO_HALAL | P52268 haliaetetus |
| 33 | 43 | 30.7 | 54 | 1 IOVO_HALIN | P05577 halliaetur 1 |

| | | | | | |
|----|----|------|-----|--------------|--------------------|
| 34 | 43 | 30.7 | 55 | 1 COX4_SAIUS | O46590 saimiri ust |
| 35 | 43 | 30.7 | 124 | 1 COX4_SAIUS | O46592 saimiri sci |
| 36 | 43 | 30.7 | 150 | 1 VGD_BPALK3 | P08765 bacterioph |
| 37 | 43 | 30.7 | 150 | 1 VGD_BPPHK | Q38039 bacterioph |
| 38 | 43 | 30.7 | 151 | 1 VGD_BPPHK | P03637 bacterioph |
| 39 | 43 | 30.7 | 152 | 1 VGD_BPG4 | P03638 bacterioph |
| 40 | 43 | 30.7 | 182 | 1 APT_PSEAE | O04633 pseudomonas |
| 41 | 43 | 30.7 | 379 | 1 YGT5_YEAST | P53100 saccharomyc |
| 42 | 43 | 30.7 | 428 | 1 EPC_HUMAN | P01854 homo sapien |
| 43 | 43 | 30.7 | 492 | 1 IRK3_CHICK | Q90854 gallus gall |
| 44 | 43 | 30.7 | 501 | 1 IRK3_HUMAN | P48549 homo sapien |
| 45 | 43 | 30.7 | 501 | 1 IRK3_MOUSE | P35562 mus musculu |

ALIGNMENTS

RESULT 1

EPC_MOUSE
ID EPC_MOUSE STANDARD; PRT; 421 AA.
AC P06336; P01856;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG EPSILON CHAIN C REGION.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84236092; PubMed=6329728;
RA Ishida N., Ueda S., Hayashida H., Miyata T., Honjo T.;
RT "The nucleotide sequence of the mouse immunoglobulin epsilon gene:
RT comparison with the human epsilon gene sequence.";
RL EMBO J. 1:1117-1123(1982).
RN [2]
RP REVISIONS.
RA Honjo T.;
RN Submitted (APR-1986) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 34-421 FROM N.A.
RX MEDLINE=83117774; PubMed=6818553;
RA Liu F.-T., Albrandt K., Sutcliffe J.G., Katz D.H.;
RT "Cloning and nucleotide sequence of mouse immunoglobulin epsilon
RT chain cDNA";
RL Proc. Natl. Acad. Sci. U.S.A. 79:7852-7856(1982).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; X01857; CAA25977.1; -;
DR EMBL; X01857; CAA25978.1; -;
DR PIR; A02145; EHMSS.
DR PIR; A02144; EHMSS.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 4.
DR SMART; SM00407; IGcl; 2.
DR SMART; SM00410; IG_like; 2.
DR PROSITE; PS00290; IG_MHC; 3.
DR Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1 1
FT DOMAIN 1 90 CH1.
FT DOMAIN 91 197 CH2.
FT DOMAIN 198 304 CH3.
FT DOMAIN 305 421 CH4.


```
RESULT 4
MK32_YEAST STANDARD; PRT; 363 AA.
AC P23060;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MAK32 PROTEIN.
GN MAK32 OR YCR019W OR YCR19W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89131254; PubMed=3916862;
RA Toh-E A., Sahashi Y.;
RT "The PET18 locus of Saccharomyces cerevisiae: a complex locus
containing multiple genes.";
RL Yeast 1:159-171(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Feldmann H., Mannhaupt G., Vetter I.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: NECESSARY FOR THE STRUCTURAL STABILITY OF L-A DOUBLE-
STRANDED RNA-CONTAINING PARTICLES. NECESSARY FOR GROWTH AT 37
DEGREES CELSIUS AS WELL AS FOR MAINTENANCE OF THE KILLER PLASMID.
CC -!- SIMILARITY: TO S.POMBE SPAC4G8.14C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DE EMBL; X59720; CAA42310.1; -
DR PIR; S19429; BVBYK2.
DR SGD; S0000612; MAK32.
FT CONFLICT 15 15 I -> II (IN REF. 1).
FT CONFLICT 82 82 MISSING (IN REF. 1).
SQ SEQUENCE 363 AA; 40783 MW; CFB358F8ACF6EA4C CRC64;

Query Match 34.6%; Score 48.5; DB 1; Length 363;
Best Local Similarity 55.0%; Pred. No. 5.4;
Matches 11; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 4 GYOSIVDR-PDFPKPIVRSI 22
I : |||| |||| :|
DB 61 GLKWIVDRGSDFPKVIRESI 80

RESULT 5
COX4_PITPI STANDARD; PRT; 144 AA.
AC O46585;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE IV (EC 1.9.3.1) (FRAGMENT).
COX4.
GN Pithecia pithecia (White-faced saki).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Pitheciinae;
OC Pithecia.
OX NCBI_TaxID=43777;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97277139; PubMed=9115172;
RA Wu W., Goodman M., Lomax M.I., Grossman L.I.;
RT "Molecular evolution of cytochrome c oxidase subunit IV: evidence for
```

```
RT positive selection in simian primates.";
RL J. Mol. Evol. 44:477-491(1997).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O + 4
FERRICYTOCHROME C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DE EMBL; AF042770; AAB97849.1; -
DR EMBL; AF042768; AAB97849.1; JOINED.
DR EMBL; AF042769; AAB97849.1; JOINED.
KW Oxidoreductase; Mitochondrion.
FT NON_TER 1
SQ SEQUENCE 144 AA; 16917 MW; 5573133B773E5C89 CRC64;

Query Match 34.3%; Score 48; DB 1; Length 144;
Best Local Similarity 50.0%; Pred. No. 2.4;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 3 YGQSVIVDRPDFPKPIVRSI 22
I : |||| |||| :|
DB 8 YTLQSVVDRRDYPLPDVAHV 27

RESULT 6
COX4_MOUSE STANDARD; PRT; 169 AA.
AC P19783;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE IV PRECURSOR (EC 1.9.3.1).
COX4.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Heart;
RX MEDLINE=91057158; PubMed=2173832;
RA Grossman L.I., Akamatsu M.;
RT "Nucleotide sequence of a mouse cDNA for subunit IV of cytochrome c
oxidase.";
RL Nucleic Acids Res. 18:6454-6454(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91378465; PubMed=1654830;
RA Carter R.S., Avadhani N.G.;
RT "Cloning and characterization of the mouse cytochrome c oxidase
subunit IV gene.";
RL Arch. Biochem. Biophys. 288:97-106(1991).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +
4 FERRICYTOCHROME C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M37831; AAB02139.1; -.
DR EMBL; M58034; AAB02139.1; JOINED.
DR EMBL; M37829; AAB02140.1; -.
DR EMBL; X54691; CAA38507.1; -.
DR PIR; S12142; S12142.
DR PIR; S12508; S12508.
DR PIR; S16114; S16114.
DR HSP; P00423; LOCC.
DR MGD; MGI:88473; COX4.
KW Oxidoreductase; Mitochondrion; Translt peptide.
FT TRANSIT 1 22 MITOCHONDRION.
FT CHAIN 23 169 CYTOCHROME C OXIDASE POLYPEPTIDE IV.
FT CONFLICT 34 34 A -> R (IN REF. 1).
FT CONFLICT 71 71 S -> N (IN REF. 1).
FT CONFLICT 140 141 MQ -> IE (IN REF. 1; AAB02139 AND IN REF. 2).
FT SEQUENCE 169 AA; 19530 MW; D30B1DBBE14FDBEA CRC64;

Query Match 34.3%; Score 48; DB 1; Length 169;
Best Local Similarity 36.4%; Pred. No. 2.9;
Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 3 YGQSVDRDPDPKPIVRSITL 24
I : : I I I I I : I
Db 33 YAFPTYADRRDYPDPVAHVTM 54

RESULT 7
PTA_THEME
ID PTA_THEME STANDARD; PRT; 294 AA.
AC Q9X0L4;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PHOSPHATE ACETYLTRANSFERASE (EC 2.3.1.8) (PHOSPHOTRANSACETYLASE).
GN PTA OR TM1130.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
Raft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima.";
RL Nature 399:323-329(1999).
RN [2]
RP SEQUENCE OF 1-39, AND CHARACTERIZATION.
RX MEDLINE=99173910; PubMed=10074080;
RA Bock A.-K., Glasenacher J., Schmidt R., Schoenheit P.;
RT "Purification and characterization of two extremely thermostable
enzymes, phosphate acetyltransferase and acetate kinase, from the
hyperthermophilic eubacterium Thermotoga maritima.";
RL J. Bacteriol. 181:1861-1867(1999).
CC -1- FUNCTION: IN ADDITION TO ACETYL-COA (100%), THE ENZYME ACCEPTS
CC PROPIONYL-COA (60%) AND BUTYRYL-COA (30%). HAS A TEMPERATURE
CC OPTIMUM AT 90 DEGREES CELSIUS.
CC -1- CATALYTIC ACTIVITY: ACETYL-COA + ORTHOPHOSPHATE = COA +
CC ACETYL PHOSPHATE.
CC -1- PATHWAY: LAST OF TWO STEPS IN THE CONVERSION OF ACETATE TO
CC ACETYL-COA.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE PHOSPHATE ACETYLTRANSFERASE AND
```

```
CC BUTYRYLTRANSFERASE FAMILY. MORE SIMILAR TO PTB THAN PTA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001770; AAD36206.1; -.
DR TIGR; TM1130; -.
DR InterPro; IPR002505; PTA_PTB.
DR Pfam; PF01515; PTA_PTB; 1.
KW Transferase; Acyltransferase; Complete proteome.
FT CONFLICT 11 11
FT CONFLICT 294 AA; 32093 MW; F4B98B3CAEI20AFB CRC64;
FT SEQUENCE 294 AA; 32093 MW; F4B98B3CAEI20AFB CRC64;

Query Match 34.3%; Score 48; DB 1; Length 294;
Best Local Similarity 46.7%; Pred. No. 5.1;
Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Qy 7 SIYDRDPDPKPIVRS 21
I : : I I I I I : I
Db 122 SVMIEDPDPRLIIS 136

RESULT 8
COX4_RAT
ID COX4_RAT STANDARD; PRT; 169 AA.
AC P10888;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE IV PRECURSOR (EC 1.9.3.1).
GN COX4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
RX MEDLINE=89240039; PubMed=2541414;
RA Goto Y., Amuro N., Okazaki T.;
RT "Nucleotide sequence of cDNA for rat brain and liver cytochrome c
oxidase subunit IV.";
RL Nucleic Acids Res. 17:2851-2851(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=WISTAR; TISSUE=Liver;
RX MEDLINE=89296488; PubMed=2544859;
RA Gopalan G., Droste M., Kadenbach B.;
RT "Nucleotide sequence of cDNA encoding subunit IV of cytochrome c
oxidase from fetal rat liver.";
RL Nucleic Acids Res. 17:4376-4376(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90237079; PubMed=2159010;
RA Yamada M., Amuro N., Goto Y., Okazaki T.;
RT "Structural organization of the rat cytochrome c oxidase subunit IV
gene.";
RL J. Biol. Chem. 265:7687-7692(1990).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=90326528; PubMed=2165254;
RA Amuro N., Yamada M., Goto Y., Okazaki T.;
RT "Complete nucleotide sequence of the gene encoding rat cytochrome c
oxidase subunit IV.";
RL Nucleic Acids Res. 18:3992-3992(1990).
RN [5]
RP SEQUENCE FROM N.A.
```

```

RX MEDLINE=91067442; PubMed=2174541;
RA Virbasius J.V., Scarpulla R.C.;
RT "The rat cytochrome c oxidase subunit IV gene family: tissue-specific
RT and hormonal differences in subunit IV and cytochrome c mRNA
RT expression.";
RL Nucleic Acids Res. 18:6581-6586(1990).
RN [6]
RP SEQUENCE OF 23-32.
RC STRAIN=WISTAR; TISSUE=Liver;
RX MEDLINE=95324529; PubMed=7601105;
RA Schaeffer H., Noack H., Halanek W., Brandt U., von Jagow G.;
RT "Cytochrome-c oxidase in developing rat heart. Enzymic properties and
RT amino-terminal sequences suggest identity of the fetal heart and the
RT adult liver isoform.";
RL Eur. J. Biochem. 230:235-241(1995).
CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X14209; CAA32426.1; -
DR EMBL; X15029; CAA33133.1; -
DR EMBL; J05423; AAA04949.1; -
DR EMBL; X54081; CAA38018.1; -
DR PIR; A35209; A35209.
DR PIR; S04070; S04070.
DR PIR; S12724; S12724.
DR PIR; S14190; S14190.
DR HSP; P00423; IOCC.
KW Oxidoreductase; Mitochondrion; Transit peptide.
FT TRANSIT 1 22 MITOCHONDRION.
FT CHAIN 23 169 CYTOCHROME C OXIDASE POLYPEPTIDE IV.
SQ SEQUENCE 169 AA; 19514 MW; F0187C76B7A1A9FE CRC64;

Query Match 33.6%; Score 47; DB 1; Length 169;
Best Local Similarity 45.5%; Pred.No. 4.1;
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 3 YGYQSIQVDRDPFKPKIVRSITL 24
| | | | | | | | | |
Db 33 YALPSYVDRRDYPLPDVAHVKL 54

RESULT 9
HEM3_CHLTR
ID HEM3_CHLTR STANDARD; PRT; 241 AA.
AC O84301;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE PORPHOBILINOGEN DEAMINASE (EC 4.3.1.8) (PBG)
DE (HYDROXYMETHYLBILANE SYNTHASE) (HMB) (PRE-UROPORPHYRINOGEN SYNTHASE).
GN HEMC OR CT299.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/UW-3/CX;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;

```

```

RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
CC -1- FUNCTION: TETRAPOLYMERIZATION OF THE MONOPYRROLE PBG INTO THE
CC HYDROXYMETHYLBILANE PREUROPORPHYRINOGEN IN SEVERAL DISCRETE STEPS
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 4 PORPHOBILINOGEN + H(2)O ->
CC HYDROXYMETHYLBILANE + 4 NH(3).
CC -1- PATHWAY: FOURTH STEP IN PORPHYRIN BIOSYNTHESIS BY THE C5 PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE HMB5 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001302; AAC67892.1; -
DR HSP; P05983; IAH5.
DR InterPro; IPR000860; Porphobil_deam.
DR Pfam; PF01379; Porphobil_deam.1.
DR PROSITE; PS00533; PORPHOBILINOGEN_DEAM; FALSE_NEG.
KW Porphyrin biosynthesis; Lyase; Complete proteome.
SQ SEQUENCE 241 AA; 26996 MW; F89FE00C20BF7044 CRC64;

Query Match 33.6%; Score 47; DB 1; Length 241;
Best Local Similarity 47.8%; Pred.No. 5.9;
Matches 11; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 1 CGYGYQSIQVDRDPFKPKIVRSIT 23
| | | | | | | | | |
Db 90 CDLGIHSKADLPENPKATVVSIT 112

RESULT 10
KPCL_LYTPI
ID KPCL_LYTPI STANDARD; PRT; 658 AA.
AC Q25378;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PROTEIN KINASE C (EC 2.7.1.1).
GN PKC1.
OS Lytechinus pictus (Painted sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Temnopleuroidea; Toxopneustidae;
OC Lytechinus.
OX NCBI_TaxID=7653;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Ovary;
RA Rakow T.L., Shen S.S.;
RT "Molecular cloning and characterization of protein kinase C from the
RT sea urchin Lytechinus pictus.";
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS IS A CALCIUM-ACTIVATED, PHOSPHOLIPID-DEPENDENT,
CC SERINE- AND THREONINE-SPECIFIC ENZYME (BY SIMILARITY).
CC -1- FUNCTION: PKC IS ACTIVATED BY DIACYLGLYCEROL WHICH IN TURN
CC PHOSPHORYLATES A RANGE OF CELLULAR PROTEINS. PKC ALSO SERVES AS
CC THE RECEPTOR FOR PHORBOL ESTERS, A CLASS OF TUMOR PROMOTERS (BY
CC SIMILARITY).
CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
CC BINDING DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 C2 DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC PKC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
```

DR EMBL; U02967; AAA03447.1; -

DR InterPro; IPR000008; C2.

DR InterPro; IPR002219; DAG_PE-bind.

DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR000961; Pkinase_C.

DR InterPro; IPR002290; Ser_thr_kin_actsite.

DR InterPro; IPR001245; Tyr_Kin.

DR Pfam; PF00168; C2; 1.

DR Pfam; PF00130; DAG_PE-bind; 2.

DR Pfam; PF00069; pkinase; 1.

DR Pfam; PF00433; pkinase_C; 1.

DR PRINTS; PR00008; DAGPEDOMAIN.

DR PRINTS; PR00109; TYRKINASE.

DR PRINTS; PR00360; C2DOMAIN.

DR SMART; SM00109; C1; 2.

DR SMART; SM00239; C2; 1.

DR SMART; SM00220; S_TKc; 1.

DR SMART; SM00133; S_TK_X; 1.

DR PROSITE; PS00499; C2_DOMAIN_1; 1.

DR PROSITE; PS00004; C2_DOMAIN_2; 1.

DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.

DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 2.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

KW ATP-binding; Transferase; Serine/threonine-protein kinase;

KW Phorbol-ester binding; Zinc; Repeat.

FT DOMAIN 28 77 PHORBOL-ESTER AND DAG BINDING 1.

FT DOMAIN 93 142 PHORBOL-ESTER AND DAG BINDING 2.

FT DOMAIN 164 251 C2 DOMAIN.

FT DOMAIN 325 583 PROTEIN KINASE.

FT NP_BIND 331 339 ATP (BY SIMILARITY).

FT BINDING 354 354 ATP (BY SIMILARITY).

FT ACT_SITE 449 449 BY SIMILARITY.

SQ SEQUENCE 658 AA; 74871 MW; 74B5A27A49C835A2 CRC64;

Query Match 32.9%; Score 46; DB 1; Length 658;

Best Local Similarity 43.5%; Pred. No. 24;

Matches 10; Conservative 7; Mismatches 4; Indels 2; Gaps 2;

Oy 5 YQ5IVDR-PDFPKPIVR-SITLC 25

Db 533 FQ5IMEHVPSPKMSRESVTMC 555

RESULT 11

ID PURL_PYRHO STANDARD; PRT; 705 AA.

AC O59621;

DT 20-AUG-2001 (Rel. 40, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DE PHOSPHORIBOSYLFORMYLGLYCINAMIDINE SYNTHASE II (EC 6.3.5.3) (FGAM SYNTHASE II).

GN PURL OR PH1953.

OS Pyrococcus horikoshii.

OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.

OX NCBI_TaxID=53953;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OT3;

RX MEDLINE=98344137; PubMed=9679194;

RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y., Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y., Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y., Funahashi Y., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,

Masuchi Y., Shizuya H., Kikuchi H.;

Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3.;

RL DNA Res. 5:55-76(1998).

CC -!- CATALYTIC ACTIVITY: ATP + 5'-PHOSPHORIBOSYLFORMYLGLYCINAMIDE + L-GLUTAMINE + H(2)O = ADP + ORTHOPHOSPHATE + 5'-PHOSPHORIBOSYL-FORMYLGLYCINAMIDINE + L-GLUTAMATE.

CC -!- PATHWAY: DE NOVO PURINE BIOSYNTHESIS; FOURTH STEP.

CC -!- SUBUNIT: HETERODIMER OF TWO SUBUNITS, PURQ AND PURL.

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE FGAMS FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AP000007; BAA31080.1; -

DR InterPro; IPR000728; AIRS_related.

DR Pfam; PF00586; AIRS; 2.

KW Purine biosynthesis; Ligase; ATP-binding; Complete proteome.

FT NP_BIND 89 100

SQ SEQUENCE 705 AA; 78540 MW; 57550733AAA4DE29 CRC64;

Query Match 32.9%; Score 46; DB 1; Length 705;

Best Local Similarity 44.4%; Pred. No. 26;

Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Oy 5 YQ5IVDRPDFPKPIVRSI 22

Db 513 YNEIVDRPIKPTPVAGV 530

RESULT 12

APT_SYNY3

ID APT_SYNY3 STANDARD; PRT; 172 AA.

AC P73935;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE ADENINE PHOSPHORIBOSYLTRANSFERASE (EC 2.4.2.7) (APRT).

GN APT OR SLL1430.

OS Synechocystis sp. (strain PCC 6803).

OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

OX NCBI_TaxID=1148;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97061201; PubMed=8905231;

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hirose M., Sugita M., Sugita M., Sasamoto S., Kimura T., Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M., Tabata S.;

RT *Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.*;

RL DNA Res. 3:109-136(1996).

CC -!- FUNCTION: CATALYSES A SALVAGE REACTION RESULTING IN THE FORMATION OF AMP, THAT IS ENERGICALLY LESS COSTLY THAN DE NOVO SYNTHESIS.

CC -!- CATALYTIC ACTIVITY: AMP + PYROPHOSPHATE = ADENINE + 5-PHOSPHO-ALPHA-D-RIBOSE 1-DIPHOSPHATE.

CC -!- PATHWAY: PURINE SALVAGE.

CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -!- SIMILARITY: BELONGS TO THE PURINE/PYRIMIDINE PHOSPHORIBOSYLTRANSFERASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

[5]
SEQUENCE OF 1-12.
RC STRAIN-K12 / EMG2; PubMed=9298646;
RX MEDLINE-97443975;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
in the genome of *Escherichia coli* K-12.";
RL Electrophoresis 18:1259-1313(1997).
CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
TO BE UBIQUINONE. DOES COUPLE THE REDOX REACTION TO PROTON
TRANSLLOCATION AND THUS CONSERVES THE REDOX ENERGY IN A PROTON
GRADIENT.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -1- COFACTOR: MAY BIND TWO 4FE-4S CLUSTER AND ONE 2FE-2S CLUSTER.
CC -1- SUBUNIT: COMPOSED OF 13 DIFFERENT SUBUNITS. SUBUNIT NUOCD, E,
F, AND, G CONSTITUTE THE PERIPHERAL SECTOR OF THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I 75 KDA SUBUNIT FAMILY.
CC -1- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
FRAMESHIFTS IN POSITIONS 714 AND 805.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; AE000317; AAC75343.1; ALT_INIT.
DR EMBL; D90859; BAA16111.1; ALT_INIT.
DR EMBL; D90860; BAA16116.1; ALT_INIT.
DR EMBL; X68301; CAA48366.1; ALT_FRAME.
DR EMBL; L25055; AAA03538.1;
DR PIR; S37064; S37064.
DR PIR; S38316; S38316.
DR Ecogene; EG12087; nuog.
DR InterPro; IPR000283; Complex1_75k.
DR InterPro; IPR001041; Ferredoxin.
DR Pfam; PF00111; fer2; 1.
DR PROSITE; PS00641; COMPLEX1_75K_1; 1.
DR PROSITE; PS00642; COMPLEX1_75K_2; 1.
DR PROSITE; PS00643; COMPLEX1_75K_3; 1.
DR Oxidoreductase; NAD; Ubiquinone; Iron-sulfur; 4Fe-4S;
Complete proteome.
KW INT_MET 0
KW METAL 22 22 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 33 33 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 44 44 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 47 47 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 102 102 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 105 105 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 111 111 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 150 150 IRON-SULFUR (4FE-4S) (POTENTIAL).
KW METAL 153 153 IRON-SULFUR (4FE-4S) (POTENTIAL).
KW METAL 200 200 IRON-SULFUR (4FE-4S) (POTENTIAL).
KW METAL 187 187 T -> Q (IN REF. 3).
KW CONFLICT 209 209 T -> K (IN REF. 3).
KW CONFLICT 389 389 MISSING (IN REF. 3).
KW CONFLICT 647 647 S -> T (IN REF. 3).
SQ SEQUENCE 907 AA; 100168 MW; 4C258FEEF36533F2 CRC64;

Query Match 32.18; Score 45; DB 1; Length 907;
Best Local Similarity 57.18; Pred. No. 49;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 2 GYGQSIIVDRPDPF 15
|||||::|||
Db 268 GYGYNLKDPRQP 281

RESULT 15
NUOG_SALTY
ID NUOG_SALTY STANDARD; PRT; 907 AA.
AC P33900;
DT 01-FEB-1994 (Rel. 28, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NADH DEHYDROGENASE I CHAIN G (EC 1.6.5.3) (NADH-UBIQUINONE
OXIDOREDUCTASE CHAIN G) (NUO7).
GN NUOG.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RN SEQUENCE OF 1-611 FROM N.A.
RX MEDLINE-94052195; PubMed=8234329;
RA Archer C.D., Wang X., Elliott T.;
RT "Mutants defective in the energy-conserving NADH dehydrogenase of
Salmonella typhimurium identified by a decrease in energy-dependent
proteolysis after carbon starvation.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:9877-9881(1993).
[2]
RN SEQUENCE OF 611-907 FROM N.A.
RC STRAIN-F98;
RX MEDLINE-98037521; PubMed=9371470;
RA Zhang-Barber L.Z., Turner A.K., Martin G., Fraenkel G., Dougan G.,
Barrow P.A.;
RT "Influence of genes encoding proton-translocating enzymes on
suppression of Salmonella typhimurium growth and colonization.";
RL J. Bacteriol. 179:7186-7190(1997).
CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
TO BE UBIQUINONE. DOES COUPLE THE REDOX REACTION TO PROTON
TRANSLLOCATION AND THUS CONSERVES THE REDOX ENERGY IN A PROTON
GRADIENT.
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -1- COFACTOR: MAY BIND TWO 4FE-4S CLUSTER AND ONE 2FE-2S CLUSTER.
CC -1- SUBUNIT: COMPOSED OF 13 DIFFERENT SUBUNITS. SUBUNIT NUOCD, E,
F, AND, G CONSTITUTE THE PERIPHERAL SECTOR OF THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I 75 KDA SUBUNIT FAMILY.
CC -1- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
FRAMESHIFTS IN POSITIONS 714 AND 805.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; L22504; AAA16063.1; ALT_INIT.
DR EMBL; L42521; -; NOT_ANNOTATED_CDS.
DR StyGene; SG10257; nuog.
DR InterPro; IPR000283; Complex1_75k.
DR InterPro; IPR001041; Ferredoxin.
DR Pfam; PF00111; fer2; 1.
DR PROSITE; PS00641; COMPLEX1_75K_1; 1.
DR PROSITE; PS00642; COMPLEX1_75K_2; 1.
DR PROSITE; PS00643; COMPLEX1_75K_3; 1.
KW Oxidoreductase; NAD; Ubiquinone; Iron-sulfur; 4Fe-4S.
KW INT_MET 0
KW BY SIMILARITY
KW METAL 22 22 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 33 33 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 44 44 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 47 47 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 102 102 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 105 105 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 111 111 IRON-SULFUR (2FE-2S) (POTENTIAL).
KW METAL 150 150 IRON-SULFUR (4FE-4S) (POTENTIAL).
KW METAL 153 153 IRON-SULFUR (4FE-4S) (POTENTIAL).
KW METAL 156 156 IRON-SULFUR (4FE-4S) (POTENTIAL).

FT METAL 200 200 IRON-SULFUR (4FE-4S) (POTENTIAL).
SQ SEQUENCE 907 AA; 100000 MW; F636867F182BAC73 CRC64;

Query Match 32.1%; Score 45; DB 1; Length 907;
Best Local Similarity 57.1%; Pred. No. 49;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Oy 2 GYGYSIVDRPDP 15
|||| : ||| |
Db 268 GYGYNLKDPRQP 281

Search completed: March 4, 2002, 13:08:43
Job time: 640 sec

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:08:12 ; Search time 79.15 Seconds
(without alignments)
46.201 Million cell updates/sec

Title: US-09-701-623C-8
Perfect score: 140
Sequence: 1 CGYGYQSVDRDPFKPIVRITLC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL17.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 51.5 | 36.8 | 186 | 5 Q27123 | Q27123 urechis cau |
| 2 | 49 | 35.0 | 165 | 1 Q93701 | Q93701 sulfolobus |
| 3 | 49 | 35.0 | 433 | 2 Q9RTJ1 | Q9RTJ1 deinococcus |
| 4 | 49 | 35.0 | 705 | 1 Q9UXW6 | Q9UXW6 pyrococcus |
| 5 | 49 | 35.0 | 1449 | 5 Q9V917 | Q9V917 drosophila |
| 6 | 48.5 | 34.6 | 722 | 13 Q9PSZ5 | Q9PSZ5 lampetra ja |
| 7 | 48 | 34.3 | 977 | 13 Q91925 | Q91925 xenopus lae |
| 8 | 48 | 34.3 | 1259 | 10 Q40983 | Q40983 pisum sativ |
| 9 | 47 | 33.6 | 534 | 12 Q65856 | Q65856 beet yellow |
| 10 | 46.5 | 33.2 | 182 | 5 Q9V108 | Q9V108 drosophila |
| 11 | 46.5 | 33.2 | 210 | 2 Q9RUH0 | Q9RUH0 deinococcus |
| 12 | 46.5 | 33.2 | 711 | 10 Q9RG55 | Q9RG55 arabidopsis |
| 13 | 46 | 32.9 | 351 | 2 Q69727 | Q69727 mycobacteri |
| 14 | 46 | 32.9 | 376 | 5 Q94259 | Q94259 caenorhabdi |
| 15 | 46 | 32.9 | 705 | 1 Q59621 | Q59621 pyrococcus |
| 16 | 46 | 32.9 | 715 | 3 Q12204 | Q12204 saccharomyc |
| 17 | 45 | 32.1 | 525 | 4 Q9NPX6 | Q9NPX6 homo sapien |
| 18 | 45 | 32.1 | 1354 | 13 Q9W6B2 | Q9W6B2 xenopus lae |
| 19 | 44.5 | 31.8 | 308 | 2 Q9AN61 | Q9AN61 bradyrhizob |

| | | | | | | |
|----|------|------|------|----|--------|--------------------|
| 20 | 44.5 | 31.8 | 330 | 2 | Q9CB84 | Q9CB84 mycobacteri |
| 21 | 44.5 | 31.8 | 366 | 10 | Q9ATS0 | Q9ATS0 bothriochlo |
| 22 | 44.5 | 31.8 | 392 | 5 | Q9VY83 | Q9VY83 drosophila |
| 23 | 44.5 | 31.8 | 467 | 9 | Q9AZ43 | Q9AZ43 bacterioph |
| 24 | 44.5 | 31.8 | 757 | 5 | Q9VNF8 | Q9VNF8 drosophila |
| 25 | 44.5 | 31.8 | 769 | 5 | Q9NFN9 | Q9NFN9 drosophila |
| 26 | 44 | 31.4 | 46 | 12 | Q88403 | Q88403 simian t-ce |
| 27 | 44 | 31.4 | 116 | 2 | Q9L4B7 | Q9L4B7 uncultured |
| 28 | 44 | 31.4 | 116 | 2 | Q9L4A5 | Q9L4A5 uncultured |
| 29 | 44 | 31.4 | 182 | 11 | Q9R1T0 | Q9R1T0 mus musculu |
| 30 | 44 | 31.4 | 222 | 4 | Q9BUX1 | Q9BUX1 homo sapien |
| 31 | 44 | 31.4 | 699 | 1 | Q9YFR1 | Q9YFR1 aeropyrum p |
| 32 | 44 | 31.4 | 706 | 2 | Q9HWD2 | Q9HWD2 pseudomonas |
| 33 | 44 | 31.4 | 767 | 5 | Q15882 | Q15882 trypanosoma |
| 34 | 44 | 31.4 | 963 | 2 | Q9XCF5 | Q9XCF5 mycobacteri |
| 35 | 44 | 31.4 | 1213 | 5 | Q9VE06 | Q9VE06 drosophila |
| 36 | 44 | 31.4 | 1219 | 5 | Q9YLA8 | Q9YLA8 drosophila |
| 37 | 43.5 | 31.1 | 143 | 2 | Q9XOG9 | Q9XOG9 thermotoga |
| 38 | 43.5 | 31.1 | 257 | 2 | P72628 | P72628 synechocyst |
| 39 | 43.5 | 31.1 | 305 | 5 | Q10667 | Q10667 caenorhabdi |
| 40 | 43.5 | 31.1 | 332 | 1 | Q9YCL2 | Q9YCL2 aeropyrum p |
| 41 | 43 | 30.7 | 152 | 9 | Q9G089 | Q9G089 bacterioph |
| 42 | 43 | 30.7 | 195 | 10 | Q9SLI4 | Q9SLI4 arabidopsis |
| 43 | 43 | 30.7 | 209 | 2 | Q55037 | Q55037 synechococc |
| 44 | 43 | 30.7 | 217 | 2 | Q9EXC0 | Q9EXC0 neisseria m |
| 45 | 43 | 30.7 | 265 | 1 | Q9YC44 | Q9YC44 aeropyrum p |

ALIGNMENTS

RESULT 1

Q27123 ID Q27123 PRELIMINARY; PRT; 186 AA.
AC Q27123;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT IV.
OS Urechis caupo (Innkeeper worm) (Spoonworm).
OC Eukaryota; Metazoa; Echiura; Xenopneusta; Urechidae; Urechis.
OX NCBI_TaxID=6431;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94155469; PubMed=8111976;
RA rosenthal E.;
RT "Sequence analysis of translationally controlled maternal mRNAs from
RT_Urechis caupo.";
RL Dev. Genet. 14:485-491(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA rosenthal E.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U30468; AAA74396.1; ;
SQ SEQUENCE 186 AA; 21084 MW; CC8EC1FA5F84D766 CRC64;

Query Match 36.8%; Score 51.5; DB 5; Length 186;
Best Local Similarity 45.5%; Pred. No. 4.6;
Matches 10; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

QY 2 GYGYQSI---VDRPDFPKPIVR 20

||| : ||| |||

Db 48 GYGMNGLPTYIDLPFPAPAVR 69

RESULT 2

Q93701 ID Q93701 PRELIMINARY; PRT; 165 AA.
AC Q93701;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)

DE HYPOTHETICAL 19.2 KDA PROTEIN.
OS Sulfolobus sp.
OG Plasmid pNOB8.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=2288;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOB8H2;
RA MEDLINE=99044580; PubMed=9827331;
RX She O., Phan H., Garrett R.A., Alberts S.V., Stedman K.M., Zillig W.;
RT "Genetic profile of pNOB8 from Sulfolobus: the first conjugative
plasmid from an archaeon."
RL Extremophiles 2:417-425(1998).
RW EMBL; AJ010405; CAA09149.1; -
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 165 AA; 19189 MW; 22E6A92532874A80 CRC64;

Query Match 35.0%; Score 49; DB 1; Length 165;
Best Local Similarity 58.8%; Pred. No. 9.7;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 5 YQSIIVDRPDPFKPIVRSI 21
Db 71 YERIDKDPNRPPIVRSI 87

RESULT 3
Q9RTJ1
ID Q9RTJ1 PRELIMINARY; PRT; 433 AA.
AC Q9RTJ1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CONSERVED HYPOTHETICAL PROTEIN.
GN DR1773.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vanathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1."
RL Science 286:1571-1577(1999).
DR EMBL; AE002019; AAFL1328.1; -
DR TIGR; DR1773; -
DR InterPro; IPR001993; Mitoch_carrier.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN 1.
KW Complete proteome.
SQ SEQUENCE 433 AA; 46392 MW; 9089C90957FE4465 CRC64;

Query Match 35.0%; Score 49; DB 2; Length 433;
Best Local Similarity 35.3%; Pred. No. 28;
Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Qy 6 QSIIVDRPDPFKPIVRSI 22
Db 117 EAAQRDPDPELLRLQ 133

RESULT 4
Q9UXW6
ID Q9UXW6 PRELIMINARY; PRT; 705 AA.

AC Q9UXW6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PHOSPHORIBOSYLFORMYLGLYCINAMIDINE SYNTHASE II (PURL).
GN PAB1201.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=2929;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Heilig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ248288; CAB50647.1; -
DR InterPro; IPR000728; AIRS_related.
DR Pfam; PF00586; AIRS; 2.
KW Complete proteome.
SQ SEQUENCE 705 AA; 77866 MW; 8D75A4AEEB102499 CRC64;

Query Match 35.0%; Score 49; DB 1; Length 705;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 5 YQSIIVDRPDPFKPIVRSI 22
Db 513 YNEVDRPVKPTPVAGI 530

RESULT 5
Q9V917
ID Q9V917 PRELIMINARY; PRT; 1449 AA.
AC Q9V917;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CGI6868 PROTEIN.
GN CGI6868.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pieriygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananatlides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RESULT
Q40983

```

ID Q40983 PRELIMINARY; PRT; 1259 AA.
AC Q40983;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE METALLOENDOPEPTIDASE.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95365331; PubMed=7638164;
RA VanderVerre P.S., Bennett T.M., Oblong J.E., Lamppa G.K.;
RT "A chloroplast processing enzyme involved in precursor maturation
RT shares a zinc-binding motif with a recently recognized family of
RT metalloendopeptidases."
RL Proc. Natl. Acad. Sci. U.S.A. 92:7177-7181(1995).
DR EMBL: U25111; AAA81472.1; -.
DR MEROPS: M16.004; -.
DR Mendel: 24274; Pissa:3178;24274.
DR InterPro: IPR001431; Peptidase_M16.
DR Pfam: PF00675; Peptidase_M16; 1.
SQ SEQUENCE 1259 AA; 139573 MW; 5DBB33ED0D123B3F CRC64;

Query Match 34.3%; Score 48; DB 10; Length 1259;
Best Local Similarity 38.1%; Pred. No. 1.3e+02;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 5 YQSVDRDPFKPIVRSITLC 25
{ | : : | | | : | | | |
Db 79 YSSVLSQPTAPVPRQSC7C 99

RESULT 9
ID Q65856 PRELIMINARY; PRT; 534 AA.
AC Q65856;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE HELLICASE.
OS Beet yellow stunt virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
OC Closterovirus.
OX NCBI_TaxID=35290;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96266429; PubMed=8661428;
RA Karasev A.V., Nikolaeva O.V., Mushegian A.R., Lee R.F., Dawson W.O.;
RT "Organization of the 3'-terminal half of beet yellow stunt virus
RT genome and implications for the evolution of closteroviruses."
RL Virology 221:199-207(1996).
DR EMBL: U51931; AAC55663.1; -.
SQ SEQUENCE 534 AA; 60979 MW; 65F0D86E486FE6F2 CRC64;

Query Match 33.6%; Score 47; DB 12; Length 534;
Best Local Similarity 42.4%; Pred. No. 70;
Matches 14; Conservative 1; Mismatches 6; Indels 12; Gaps 2;

QY 1 CGVG-----YQSVDRPD-----FPKPIVRS 21
| | | | | | | | | | | | | |
Db 357 CYGTGTARQVRVVRKPDYELDGIFSKPIVMS 389

RESULT 10
Q9VIQ8
ID Q9VIQ8 PRELIMINARY; PRT; 182 AA.
AC Q9VIQ8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 23.0 KDA PROTEIN.

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CG10564 PROTEIN.
GN CG10564
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003664; AAF53857.1; -.
DR FlyBase: FBgn0032833; CG10664.
SQ SEQUENCE 182 AA; 20519 MW; 6D5953D8DE41217A CRC64;

Query Match 33.2%; Score 46.5; DB 5; Length 182;
Best Local Similarity 45.5%; Pred. No. 26;
Matches 10; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 2 GYGYS---IVDRDPFKPIVR 20
| | | | | | | | | |
Db 43 GYGMNGTACVADRVDPVLPVAVR 64

RESULT 11
Q9RUH0
ID Q9RUH0 PRELIMINARY; PRT; 210 AA.
AC Q9RUH0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 23.0 KDA PROTEIN.
```

QY 1 CGYGYSIV-----DRDPFKPIV 19
 ||:| |:| |:| |:|
 Db 382 CGGLSSVYKAQSSSKPGTPPVI 405

RESULT 13
 069727 PRELIMINARY; PRT; 351 AA.

AC 069727;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE HYPOTHETICAL 38.6 KDA PROTEIN.
 GN RV376IC OR MTV025.109C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garrier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skellon S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.";
 RT Nature 393:537-544(1998).
 RL ENBL; AL022121; CAA18083.1. -
 DR TubercuList; RV376IC. -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 351 AA; 38564 MW; 1031303DBAOCB30A CRC64;

Query Match 32.9%; Score 46; DB 2; Length 351;
 Best Local Similarity 38.1%; Pred. No. 63;
 Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0

QY 5 YQSIDVRDPFKPIVRSITLC 25
 |:| |:| |:| |:|
 Db 74 YRVVAALGDPVPVARTISLC 94

RESULT 14
 094259 PRELIMINARY; PRT; 376 AA.

AC 094259;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE K04A8.1 PROTEIN.
 GN K04A8.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RN None;
 RA "Genome sequence of the nematode C. elegans: a platform for
 R investigating biology. The C. elegans Sequencing Consortium.;"
 RL Science 281:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RN None;

RA Geisel C., Bradshaw H.;
RT "The sequence of C. elegans cosmid K04A8.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: U64849; AAC48047.2; -;
SQ SEQUENCE 376 AA; 43086 MW; D73B08DB0FE7D6D1 CRC64;

Query Match 32.9%; Score 46; DB 5; Length 376;
Best Local Similarity 45.0%; Pred. No. 68;
Matches 9; Conservative 4; Mismatches 5; Indels 2; Gaps 1;

Qy 3 YQSIIVDRDPFKPIVRSI 22
| | | : | | : | | |
Db 84 FGYE--ITRPTYVPVTRMI 101

RESULT 15
O59621
ID O59621 PRELIMINARY; PRT; 705 AA.
AC O59621;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 705AA LONG HYPOTHETICAL PHOSPHORIBOSYLFORMYLGLYCINAMIDINE SYNTHASE II.
GN PH1953.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL: AP000007; BAA31080.1; -;
DR InterPro: IPR000728; AIRS_related.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00586; AIRS; 2.
DR PROSITE: PS00678; WD_REPEATS_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 705 AA; 78540 MW; 57550733AAA4DE29 CRC64;

Query Match 32.9%; Score 46; DB 1; Length 705;
Best Local Similarity 44.4%; Pred. No. 1.4e+02;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 5 YQSIIVDRDPFKPIVRSI 22
| | | | | | | | | |
Db 513 YNEIVDRPIKPTPVVAGV 530

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:56:35 ; Search time 38.04 Seconds
(without alignments)
14.789 Million cell updates/sec

Title: US-09-701-623C-8
Perfect score: 140
Sequence: 1 CGYGQSVDRPDPFKPIVRSITL 25
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 116 | 82.9 | 561 | 3 | US-09-192-545-2 |
| 2 | 64 | 45.7 | 25 | 3 | US-09-100-414B-95 |
| 3 | 64 | 45.7 | 25 | 4 | US-09-303-323-95 |
| 4 | 64 | 45.7 | 42 | 3 | US-09-100-414B-98 |
| 5 | 64 | 45.7 | 42 | 3 | US-09-100-414B-99 |
| 6 | 64 | 45.7 | 42 | 3 | US-09-100-414B-100 |
| 7 | 64 | 45.7 | 42 | 4 | US-09-303-323-98 |
| 8 | 64 | 45.7 | 42 | 4 | US-09-303-323-99 |
| 9 | 64 | 45.7 | 42 | 4 | US-09-303-323-100 |
| 10 | 64 | 45.7 | 45 | 3 | US-09-100-414B-101 |
| 11 | 64 | 45.7 | 45 | 4 | US-09-303-323-101 |
| 12 | 64 | 45.7 | 46 | 3 | US-09-100-414B-96 |
| 13 | 64 | 45.7 | 46 | 3 | US-09-100-414B-97 |
| 14 | 64 | 45.7 | 46 | 4 | US-09-303-323-96 |
| 15 | 64 | 45.7 | 46 | 4 | US-09-303-323-97 |
| 16 | 64 | 45.7 | 63 | 3 | US-09-100-414B-102 |
| 17 | 64 | 45.7 | 63 | 4 | US-09-303-323-102 |
| 18 | 48 | 34.3 | 268 | 3 | US-09-187-049-9 |
| 19 | 48 | 34.3 | 1259 | 3 | US-09-187-049-13 |
| 20 | 45 | 32.1 | 426 | 1 | US-08-336-583-2 |
| 21 | 45 | 32.1 | 426 | 5 | PCT-US95-13795-2 |
| 22 | 45 | 32.1 | 699 | 5 | PCT-US94-07297-39 |
| 23 | 45 | 32.1 | 921 | 1 | US-08-396-479B-2 |
| 24 | 45 | 32.1 | 921 | 1 | US-08-818-823-2 |
| 25 | 44.5 | 31.8 | 396 | 4 | US-09-461-474-2 |
| 26 | 43 | 30.7 | 106 | 2 | US-08-232-539D-54 |
| 27 | 43 | 30.7 | 113 | 2 | US-08-232-539D-56 |

| | | | | | | |
|----|------|------|------|---|-------------------|-------------------|
| 28 | 43 | 30.7 | 501 | 1 | US-08-103-445-4 | Sequence 4, Appli |
| 29 | 43 | 30.7 | 501 | 1 | US-08-461-690B-4 | Sequence 4, Appli |
| 30 | 43 | 30.7 | 501 | 1 | US-08-486-342-2 | Sequence 2, Appli |
| 31 | 43 | 30.7 | 501 | 1 | US-08-473-092-2 | Sequence 2, Appli |
| 32 | 43 | 30.7 | 501 | 1 | US-08-614-801A-2 | Sequence 2, Appli |
| 33 | 43 | 30.7 | 501 | 1 | US-08-066-371-2 | Sequence 2, Appli |
| 34 | 43 | 30.7 | 501 | 4 | US-09-042-494-2 | Sequence 2, Appli |
| 35 | 43 | 30.7 | 1290 | 1 | US-08-470-350B-2 | Sequence 2, Appli |
| 36 | 43 | 30.7 | 1843 | 4 | US-09-413-814-50 | Sequence 50, Appl |
| 37 | 42 | 30.0 | 482 | 2 | US-08-876-874-2 | Sequence 2, Appli |
| 38 | 41.5 | 29.6 | 248 | 3 | US-08-750-145A-18 | Sequence 18, Appl |
| 39 | 41.5 | 29.6 | 248 | 3 | US-08-975-698A-22 | Sequence 22, Appl |
| 40 | 41.5 | 29.6 | 248 | 4 | US-09-417-090-22 | Sequence 22, Appl |
| 41 | 41.5 | 29.6 | 776 | 2 | US-08-671-978A-41 | Sequence 41, Appl |
| 42 | 41 | 29.3 | 1464 | 2 | US-08-231-193A-11 | Sequence 11, Appl |
| 43 | 41 | 29.3 | 1464 | 3 | US-08-486-273A-11 | Sequence 11, Appl |
| 44 | 41 | 29.3 | 1464 | 3 | US-08-940-086A-11 | Sequence 11, Appl |
| 45 | 41 | 29.3 | 1464 | 4 | US-08-436-332B-10 | Sequence 10, Appl |

ALIGNMENTS

RESULT 1
US-09-192-545-2
; Sequence 2, Application US/09192545
; Patent No. 6118044
; GENERAL INFORMATION:
; APPLICANT: Karasuyama, Hajime
; APPLICANT: Yonekawa, Hiromichi
; APPLICANT: Iaya, Choji
; TITLE OF INVENTION: Transgenic Animal Allergy Models and Methods for Their Use
; FILE REFERENCE: 799P79570
; CURRENT APPLICATION NUMBER: US/09/192,545
; CURRENT FILING DATE: 1998-11-13
; EARLIER APPLICATION NUMBER: JP HEI 9-313989
; EARLIER FILING DATE: 1997-11-14
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE: Description of Artificial Sequence: Designed heavy
US-09-192-545-2

Query Match 82.9%; Score 116; DB 3; Length 561;
Best Local Similarity 91.3%; Pred. No. 6.8e-10;
Matches 21; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GYGQSVDRPDPFKPIVRSITL 24
|||||:|||||
DB 419 GYGQCVDRPDPFKPIVRSITL 441

RESULT 2
US-09-100-414B-95
; Sequence 95, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Y1
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-95

Query Match 45.7%; Score 64; DB 3; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0019;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYQSVDRPDPPKPIVRSITLC 25
|| ||| | | : || | |
Db 1 CGETYQSRVTHPLPRALMRSTTKC 25

RESULT 3
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-303-323-95

Query Match 45.7%; Score 64; DB 4; Length 25;
Best Local Similarity 48.0%; Pred. No. 0.0019;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYQSVDRPDPPKPIVRSITLC 25
|| ||| | | : || | |
Db 1 CGETYQSRVTHPLPRALMRSTTKC 25

RESULT 4
US-09-100-414B-98
; Sequence 98, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-100-414B-98

Query Match 45.7%; Score 64; DB 3; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0034;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYQSVDRPDPPKPIVRSITLC 25
|| ||| | | : || | |
Db 18 CGETYQSRVTHPLPRALMRSTTKC 42

RESULT 5
US-09-100-414B-99
; Sequence 99, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Morgan & Finnegan, L.L.P.
;; STREET: 345 Park Avenue
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10154-0054
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC Windows
;; SOFTWARE: Word 97
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/100,414B
;; FILING DATE: 20-JUNE-1998
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Maria H. Lin
;; REGISTRATION NUMBER: 29,323
;; REFERENCE/DOCKET NUMBER: 1151-4157
;; TELEPHONE: 212-758-4800
;; TELEFAX: 212-751-6849
;; INFORMATION FOR SEQ ID NO: 99:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 42 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-100-414B-99

Query Match 45.7%; Score 64; DB 3; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0034;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYQSIVDRPDPFKPIVRSITLC 25
|| ||| | | : || | |
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 6
US-09-100-414B-100
;; Sequence 100, Application US/09100414B
;; Patent No. 6025468
;; GENERAL INFORMATION:
;; APPLICANT: Wang, Chang Yi
;; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
;; NUMBER OF SEQUENCES: 106
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Morgan & Finnegan, L.L.P.
;; STREET: 345 Park Avenue
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10154-0054
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC Windows
;; SOFTWARE: Word 97
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/100,414B
;; FILING DATE: 20-JUNE-1998
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Maria H. Lin
;; REGISTRATION NUMBER: 29,323
;; REFERENCE/DOCKET NUMBER: 1151-4157
;; TELEPHONE: 212-758-4800
;; TELEFAX: 212-751-6849

;; INFORMATION FOR SEQ ID NO: 100:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 42 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-100-414B-100

Query Match 45.7%; Score 64; DB 3; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0034;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYQSIVDRPDPFKPIVRSITLC 25
|| ||| | | : || | |
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 7
US-09-303-323-98
;; Sequence 98, Application US/09303323
;; Patent No. 6228987
;; GENERAL INFORMATION:
;; APPLICANT: Wang, Chang Yi
;; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
;; NUMBER OF SEQUENCES: 106
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Morgan & Finnegan, L.L.P.
;; STREET: 345 Park Avenue
;; CITY: New York
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10154-0054
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC Windows
;; SOFTWARE: Word 97
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/303,323
;; FILING DATE: 30-APR-1999
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 09/100,414
;; FILING DATE: 20-JUNE-1998
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Maria H. Lin
;; REGISTRATION NUMBER: 29,323
;; REFERENCE/DOCKET NUMBER: 1151-4157
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-758-4800
;; TELEFAX: 212-751-6849
;; INFORMATION FOR SEQ ID NO: 98:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 42 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-09-303-323-98

Query Match 45.7%; Score 64; DB 4; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0034;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGYQSIVDRPDPFKPIVRSITLC 25
|| ||| | | : || | |
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 8
US-09-303-323-99

Sequence 99, Application US/09303323
Patent No. 6228987
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-99

Query Match 45.7%; Score 64; DB 4; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0034;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGQSVDRDPDPKPIVRSITLC 25
|| ||| | | | : : || | |
DB 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 9
US-09-303-323-100
Sequence 100, Application US/09303323
Patent No. 6228987
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-100

Query Match 45.7%; Score 64; DB 4; Length 42;
Best Local Similarity 48.0%; Pred. No. 0.0034;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYGQSVDRDPDPKPIVRSITLC 25
|| ||| | | | : : || | |
DB 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 10
US-09-100-414B-101
Sequence 101, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-101

Query Match 45.7%; Score 64; DB 3; Length 45;

Best Local Similarity 48.0%; Pred. No. 0.0037;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYQSQIVDRPDPKPIVRSITLC 25
|| ||| | | | | | : || | |
Db 21 CGETYQSRVTHPLPALMRSTTKC 45

RESULT 11

US-09-303-323-101
; Sequence 101, Application US/09303323
; Patent No. 6228987

GENERAL INFORMATION:

; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999

CLASSIFICATION:

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414

ATTORNEY/AGENT INFORMATION:

; FILING DATE: 20-JUNE-1998
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO:

; 101:

SEQUENCE CHARACTERISTICS:

; LENGTH: 45 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-09-303-323-101

Query Match 45.7%; Score 64; DB 4; Length 45;
Best Local Similarity 48.0%; Pred. No. 0.0037;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYQSQIVDRPDPKPIVRSITLC 25
|| ||| | | | | | : || | |
Db 21 CGETYQSRVTHPLPALMRSTTKC 45

RESULT 12

US-09-100-414B-96
; Sequence 96, Application US/09100414B
; Patent No. 6025468

GENERAL INFORMATION:

; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue

; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO:

; 96:

SEQUENCE CHARACTERISTICS:

; LENGTH: 46 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-09-100-414B-96

Query Match 45.7%; Score 64; DB 3; Length 46;
Best Local Similarity 48.0%; Pred. No. 0.0038;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 CGYQSQIVDRPDPKPIVRSITLC 25
|| ||| | | | | | : || | |
Db 22 CGETYQSRVTHPLPALMRSTTKC 46

RESULT 13

US-09-100-414B-97
; Sequence 97, Application US/09100414B
; Patent No. 6025468

GENERAL INFORMATION:

; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; TITLE OF INVENTION: IMMUNOGENS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY

; COUNTRY: USA

; ZIP: 10154-0054

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849

INFORMATION FOR SEQ ID NO:

; 97:

SEQUENCE CHARACTERISTICS:

; LENGTH: 46 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-97

Query Match 45.7%; Score 64; DB 3; Length 46;
Best Local Similarity 48.0%; Pred. No. 0.0038;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CGYGYQSIQVDRPDPKPIVRSITLC 25
|| ||| | | | : || | |
Db 22 CGETYQSRVTHPLPALMRSTTKC 46

RESULT 14

US-09-303-323-96
; Sequence 96, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:

; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY

; COUNTRY: USA

; ZIP: 10154-0054

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC Windows

; SOFTWARE: Word 97

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/303,323

; FILING DATE: 30-APR-1999

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/100,414

; FILING DATE: 20-JUNE-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Maria H. Lin

; REGISTRATION NUMBER: 29,323

; REFERENCE/DOCKET NUMBER: 1151-4157

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-758-4800

; TELEFAX: 212-751-6849

; INFORMATION FOR SEQ ID NO: 96:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 46 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-09-303-323-96

Query Match 45.7%; Score 64; DB 4; Length 46;
Best Local Similarity 48.0%; Pred. No. 0.0038;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CGYGYQSIQVDRPDPKPIVRSITLC 25
|| ||| | | | : || | |
Db 22 CGETYQSRVTHPLPALMRSTTKC 46

RESULT 15

US-09-303-323-97

; Sequence 97, Application US/09303323

; Patent No. 6228987

; GENERAL INFORMATION:

; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY

; COUNTRY: USA

; ZIP: 10154-0054

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC Windows

; SOFTWARE: Word 97

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/303,323

; FILING DATE: 30-APR-1999

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/100,414

; FILING DATE: 20-JUNE-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Maria H. Lin

; REGISTRATION NUMBER: 29,323

; REFERENCE/DOCKET NUMBER: 1151-4157

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-758-4800

; TELEFAX: 212-751-6849

; INFORMATION FOR SEQ ID NO: 97:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 46 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-09-303-323-97

Query Match 45.7%; Score 64; DB 4; Length 46;
Best Local Similarity 48.0%; Pred. No. 0.0038;
Matches 12; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 CGYGYQSIQVDRPDPKPIVRSITLC 25
|| ||| | | | : || | |
Db 22 CGETYQSRVTHPLPALMRSTTKC 46

Search completed: March 4, 2002, 12:56:35
Job time: 257 sec

THIS PAGE BLANK (user)

A:Cross-references: GB:S55276; NID:g263168; PIDN:AAB24858.1; PID:g263169
A:Experimental source: B cell myeloma U-266
A>Note: sequence extracted from NCBI backbone (NCBIP:125299)
A:Accession: A46536
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 401-428 <HE3>
A:Cross-references: GB:S53497; NID:g263162; PIDN:AAB24855.1; PID:g263163
A:Experimental source: B cell myeloma U-266
A>Note: sequence extracted from NCBI backbone (NCBIP:123483)
C:Genetics:
A:Gene: GDB: IGHE
A:Cross-references: GDB:I19335; OMIM:147180
A:Map position: 14q32.33-14q32.33
A:Introns: 1/1; 104/1; 211/1; 319/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (L) chain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: alternative splicing; duplication; glycoprotein; heterotetramer; immunoglobulin F:22-87/Domain: immunoglobulin homology <IM1>
F:128-195/Domain: immunoglobulin homology <IM2>
F:232-301/Domain: immunoglobulin homology <IM3>
F:338-407/Domain: immunoglobulin homology <IM4>
F:14/Disulfide bonds: interchain (to light chain) #status predicted
F:15-105,29-85,135-193,239-299,345-405/disulfide bonds: #status predicted
F:21,49,99,146,252,275/Binding site: carbohydrate (Asn) (covalent) #status experiment
F:121,209/disulfide bonds: interchain (to heavy chain) #status predicted

Query Match 52.6%; Score 71; DB 1; Length 428;
Best Local Similarity 56.5%; Pred. No. 0.0081;
Matches 13; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREVVRSTAK 24
||||| :||| ||| :||| |
Db 294 GETQCRTVTHPLPRALMRSTTK 316

RESULT 3
I68730
IgE chain C3 region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C:Accession: I68730
R:Shinkai, Y.; Nakachi, H.; Honjo, T.; Okumura, K.
ImmunoGenetics 27, 288-292, 1988
A>Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic acid sequences of the heavy chain genes
A:Reference number: I54443; MUID:88152907
A:Accession: I68730
A>Status: preliminary; translated from GB/EMBL/DDEJ
A:Molecule type: mRNA
A:Residues: 1-107 <RES>
A:Cross-references: GB:N22933; NID:gl94464; PIDN:AAA37915.1; PID:gl94469
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:22-90/Domain: immunoglobulin homology <IM>

Query Match 48.9%; Score 66; DB 2; Length 107;
Best Local Similarity 52.2%; Pred. No. 0.01;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREVVRSTAK 24
| | | | | : | | | | |
Db 83 GYGQCIVDHPDFPKPIVRSITK 105

RESULT 4
I68726
IgE chain C3 region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C:Accession: I68726
R:Shinkai, Y.; Nakachi, H.; Honjo, T.; Okumura, K.

Immunogenetics 27, 288-292, 1988
A:Title: Mouse immunoglobulin allotypes: Multiple differences between the nucleic acid
A:Reference number: I54443; MUID:88152907
A:Accession: I68726
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-107 <RES>
A:Cross-references: GB:M22930; NID:g194455; PIDN:AAA37911.1; PID:g194460
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:22-90/Domain: immunoglobulin homology <IMM>

Query Match 48.9%; Score 66; DB 2; Length 107;
Best Local Similarity 52.2%; Pred. No. 0.01;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREVRSIAK 24

Db 83 GYGQCIVDHPDFPKPIVRSITK 105

RESULT 5

EHMS

Ig epsilon chain C region (version 1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 16-Jul-1999
C:Accession: A02144
R:Liu, F.T.; Albrandt, K.; Sutcliffe, J.G.; Katz, D.H.
Proc. Natl. Acad. Sci. U.S.A. 79, 7852-7856, 1982
A:Title: Cloning and nucleotide sequence of mouse immunoglobulin epsilon chain cDNA.
A:Reference number: A02144; MUID:83117774
A:Accession: A02144
A:Molecule type: mRNA
A:Residues: 1-388 <LIU>
A:Cross-references: GB:J00476; NID:g194875; PIDN:AAA38085.1; PID:g387220
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (kap
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:1-44/Domain: immunoglobulin homology (fragment) <IM1>
F:81-149/Domain: immunoglobulin homology <IM2>
F:186-254/Domain: immunoglobulin homology <IM3>
F:290-361/Domain: immunoglobulin homology <IM4>
F:10,51,62,133,205,228,332,382/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 48.9%; Score 66; DB 1; Length 388;
Best Local Similarity 52.2%; Pred. No. 0.041;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREVRSIAK 24

Db 247 GYGQCIVDHPDFPKPIVRSITK 269

RESULT 6

EHRT

Ig epsilon chain C region - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 16-Jul-1999
C:Accession: A93442; A90937; A02143
R:Hellman, L.; Pettersson, U.; Engstrom, A.; Karlsson, T.; Bennich, H.
Nucleic Acids Res. 10, 6041-6049, 1982
A:Title: Structure and evolution of the heavy chain from rat immunoglobulin E.
A:Reference number: A93442; MUID:83064537
A:Accession: A93442
A:Molecule type: mRNA
A:Residues: 1-429 <HEL>
A:Experimental source: strain LOU/c/Wsl, immunocytoma IR2
R:Kindsvogel, W.R.; Reddy, E.P.; Moore, J.M.; Faust Jr., C.H.
DNA 1, 335-343, 1982
A:Title: A cloned cDNA probe for rat immunoglobulin epsilon heavy chain: construction,
A:Reference number: A90937; MUID:83182019

A:Contents: myeloma IR162
A:Accession: A90937
A:Molecule type: mRNA
A:Residues: 'N',169-307,'L',309-342 <KIN>
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (c
hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:19-80/Domain: immunoglobulin homology <IM1>
F:118-186/Domain: immunoglobulin homology <IM2>
F:223-291/Domain: immunoglobulin homology <IM3>
F:327-398/Domain: immunoglobulin homology <IM4>
F:46,99,170,240,265,369,419/Binding site: carbohydrate (Asn) (covalent) #status predi

Query Match 48.9%; Score 66; DB 1; Length 429;
Best Local Similarity 52.2%; Pred. No. 0.045;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREVRSIAK 24

Db 284 GEGYQCRVDHPDFPKPIVRSITK 306

RESULT 7

S38864

Ig epsilon chain C region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 03-Nov-2000
C:Accession: S38864
R:Kipp, B.; Becker, W.; Schlaak, M.
submitted to the EMBL Data Library, November 1993
A:Description: Combination of a defined specificity and desired isotype by cloning of
A:Reference number: S38864
A:Accession: S38864
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-548 <KITP>
A:Cross-references: EMBL:Z27397; NID:g416537; PIDN:CAA81788.1; PID:g940782
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:353-421/Domain: immunoglobulin homology <IMM>

Query Match 48.9%; Score 66; DB 2; Length 548;
Best Local Similarity 52.2%; Pred. No. 0.059;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREVRSIAK 24

Db 414 GYGQCIVDHPDFPKPIVRSITK 436

RESULT 8

S43147

Ig epsilon chain - duck (fragment)
C:Species: Anas platyrhynchos (domestic duck)
C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 11-Jan-2000
C:Accession: S43147
R:Magor, K.E.; Higgins, D.A.; Middleton, D.L.; Warr, G.W.
submitted to the EMBL Data Library, March 1994
A:Description: Evidence from duck immunoglobulin genes that IgY is the common ancesto
A:Reference number: S43145
A:Accession: S43147
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-110 <MAG>
A:Cross-references: EMBL:X78355; NID:g468612; PID:g468613
C:Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 44.4%; Score 60; DB 2; Length 110;
Best Local Similarity 52.2%; Pred. No. 0.085;
Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 2 GETYKSTVSHPDLPREVVRSTAK 24
 || : || | ||| : : ||||
 Db 86 GERFCTVQHEDLPVPLCKSIK 108

RESULT 9
 B46529
 Ig Y heavy chain (7.8S) - duck
 N:Alternate names: Ig gamma chain (7.8S)
 C:Species: Anas platyrhynchos (domestic duck)
 C:Accession: B46529; #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
 R:Magor, K.E.; Warr, G.W.; Middleton, D.; Wilson, M.R.; Higgins, D.A.
 J. Immunol. 149, 2627-2633, 1992
 A:Title: Structural relationship between the two Igy of the duck, Anas platyrhynchos: m
 A:Reference number: A46529; MUID:93017865
 A:Accession: B46529
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-572 <MAG>
 A:CROSS-references: EMBL:X65219; NID:g62442; PIDN:CAA46322.1; PID:g62443
 A:Experimental source: spleen
 A:Note: sequence extracted from NCBI backbone (NCBIP:116127)
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: immunoglobulin
 F:37-120/Domain: immunoglobulin homology <IMM>

Query Match 42.2%; Score 57; DB 2; Length 572;
 Best Local Similarity 52.2%; Pred. No. 1.4;
 Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 2 GETYKSTVSHPDLPREVVRSTAK 24
 || : || | ||| : : ||||
 Db 437 GERFCTVQHEDLPVPLCKSIK 459

RESULT 10
 S70214
 LepA protein homolog - Photobacterium sp. (fragment)
 C:Species: Photobacterium sp.
 C:Date: 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 02-Feb-2001
 C:Accession: S70214
 R:Chi, E.; Bartlett, D.H.
 Mol. Microbiol. 17, 713-726, 1995
 A:Title: An rpoE-like locus controls outer membrane protein synthesis and growth at cold
 A:Reference number: S70210; MUID:96111491
 A:Accession: S70214
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-137 <CHT>
 A:CROSS-references: EMBL:L41667; NID:g777746; PIDN:AAB31927.1; PID:g777751
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995
 C:Superfamily: GTP-binding membrane protein lepa; translation elongation factor Tu homol
 C:Keywords: GTP binding; nucleotide binding; P-loop
 F:5-134/Domain: translation elongation factor Tu homology <TU>
 F:11-18/Region: nucleotide-binding motif A (P-loop)
 F:131-134/Region: GTP-binding NKXD motif

Query Match 40.0%; Score 54; DB 2; Length 137;
 Best Local Similarity 46.4%; Pred. No. 0.84;
 Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

Qy 2 GETYK-----STVSHPDLPREVVRSTAK 25
 |||| : | | | | | | | |
 Db 68 GETYQLNFIDTPGHVDFSEVSRSLAAC 95

RESULT 11
 EHM55
 Ig epsilon chain C region (version 2) - mouse

C:Species: Mus musculus (house mouse)
 C:Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 16-Aug-1996
 C:Accession: A02145
 R:Ishida, N.; Ueda, S.; Hayashida, H.; Miyata, T.; Honjo, T.
 EMBO J. 1, 1117-1123, 1982
 A:Title: The nucleotide sequence of the mouse immunoglobulin epsilon gene: comparison
 A:Reference number: A90966; MUID:84236092
 A:Accession: A02145
 A:Molecule type: DNA
 A:Residues: 1-423 <ISH>
 A:Note: the sequence was determined from the germline gene
 C:Genetics:
 A:Introns: 91/1; 199/1; 307/1
 C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (C
 hain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
 F:16-77/Domain: immunoglobulin homology <IMM1>
 F:115-183/Domain: immunoglobulin homology <IMM2>
 F:220-288/Domain: immunoglobulin homology <IMM3>
 F:325-396/Domain: immunoglobulin homology <IMM4>
 F:23-75,122-181,227-286,332-394/Disulfide bonds: #status predicted
 F:43,84,167,239,262,417/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.0%; Score 54; DB 1; Length 423;
 Best Local Similarity 47.6%; Pred. No. 2.8;
 Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 2 GETYKSTVSHPDLPREVVRSI 22
 | | : | | | : : ||||
 Db 281 GYGVCVVDRDPDPKPIVRSI 301

RESULT 12
 I64042

GTP-binding membrane protein lepa - Haemophilus influenzae (strain Rd KW20)
 C:Species: Haemophilus influenzae
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 02-Feb-2001
 C:Accession: I64042
 R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman
 , D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
 Science 269, 496-512, 1995
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.
 A:Reference number: A64000; MUID:95350630
 A:Accession: I64042
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-598 <TIGR>
 A:CROSS-references: GB:U32687; GB:I42023; NID:g1572955; PIDN:AAC21694.1; PID:g1572960
 C:Genetics:
 A:Gene: lepa
 C:Superfamily: GTP-binding membrane protein lepa; translation elongation factor Tu ho
 C:Keywords: GTP binding; membrane protein; nucleotide binding; P-loop
 F:5-134/Domain: translation elongation factor Tu homology <TU>
 F:11-18/Region: nucleotide-binding motif A (P-loop)
 F:131-134/Region: nucleotide-binding NKXD motif
 F:162-164/Region: GTP-binding SAK/L motif
 F:17,18,53,131,132,134,162/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #

Query Match 40.0%; Score 54; DB 2; Length 598;
 Best Local Similarity 46.4%; Pred. No. 4;
 Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

Qy 2 GETYK-----STVSHPDLPREVVRSTAK 25
 |||| : | | | | | | | |
 Db 68 GETYQLNFIDTPGHVDFSEVSRSLAAC 95

RESULT 13

BVECLA
GTP-binding membrane protein lepA - Escherichia coli

C;Species: Escherichia coli
C;Date: 31-Mar-1988 #sequence_revision 05-Dec-1997 #text_change 19-Jan-2001
C;Accession: H65034; A22627

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Cohen, A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617

A;Accession: H65034

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-599 <BLAT>

A;Cross-references: GB:AE000343; GB:U00096; NID:g2367139; PIDN:AAC75622.1; PID:g1788922;

A;Experimental source: strain K-12, substrain MG1655

R;March, P.E.; Inouye, M.

J. Biol. Chem. 260, 7206-7213, 1985

A;Title: Characterization of the lep operon of Escherichia coli. Identification of the P

A;Reference number: A22627; MUID:85207751

A;Accession: A22627

A;Molecule type: DNA

A;Residues: 1-52,'N',54-155,'HRRGALFSENRRWCAGRSR',156,'SGA',160,'H',162-163,'AGS',186-26

A;Note: The authors translated the codon AAC for residue 53 as Thr

C;Comment: The overproduction of this protein is lethal to E. coli. It is present in the

codes for signal peptidase.

C;Genetics:

A;Gene: lepA

A;Map position: 55 min

C;Superfamily: GTP-binding membrane protein lepA; translation elongation factor Tu homol

C;Keywords: GTP binding; membrane protein; nucleotide binding; P-loop

F;5-134/Domain: translation elongation factor Tu homology <ETU>

F;11-18/Region: nucleotide-binding motif A (P-loop)

F;131-134/Region: GTP-binding NKXD motif

F;162-164/Region: GTP-binding SAK/L motif

F;17,18,53,131,132,134,162/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

Query Match 40.0%; Score 54; DB 1; Length 599;

Best Local Similarity 46.4%; Pred. No. 4;

Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

QY 2 GETYK----STVSHPDLPREVVRSLAK 25

Db 68 GETYQLNFIDTPGHVDFSEYVSRSLAAC 95

RESULT 14

A85903

hypothetical protein lepA [imported] - Escherichia coli (strain O157:H7)

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001

C;Accession: A85903

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-599 <STO>

A;Cross-references: GB:AE005174; NID:g12516987; PIDN:AAG57685.1; GSPDB:GN00145; UWGP:238

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: lepA

C;Superfamily: GTP-binding membrane protein lepA; translation elongation factor Tu homol

Query Match 40.0%; Score 54; DB 2; Length 599;

Best Local Similarity 46.4%; Pred. No. 4;

Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:08:43 ; Search time 24.65 Seconds
(without alignments)
37.185 Million cell updates/sec

Title: US-09-701-623C-84
Perfect score: 135
Sequence: 1 CGETKSTVSHPDLPREVRSTAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------|
| 1 | 71 | 52.6 | 428 | 1 | EPC_HUMAN |
| 2 | 66 | 48.9 | 421 | 1 | EPC_MOUSE |
| 3 | 66 | 48.9 | 429 | 1 | EPC_RAT |
| 4 | 54 | 40.0 | 598 | 1 | LEPA_HAEN |
| 5 | 54 | 40.0 | 598 | 1 | LEPA_PASMU |
| 6 | 54 | 40.0 | 599 | 1 | LEPA_ECOLI |
| 7 | 54 | 40.0 | 602 | 1 | LEPA_CHLMU |
| 8 | 54 | 40.0 | 602 | 1 | LEPA_CHLTR |
| 9 | 52.5 | 38.9 | 450 | 1 | MUC_CANFA |
| 10 | 51 | 37.8 | 340 | 1 | ALC2_HUMAN |
| 11 | 51 | 37.8 | 391 | 1 | MUCB_HUMAN |
| 12 | 51 | 37.8 | 454 | 1 | MUC_HUMAN |
| 13 | 51 | 37.8 | 457 | 1 | MUC_SUNMU |
| 14 | 51 | 37.8 | 458 | 1 | MUC_RABIT |
| 15 | 51 | 37.8 | 479 | 1 | MUCM_RABIT |
| 16 | 50 | 37.0 | 600 | 1 | LEPA_AQAE |
| 17 | 50 | 37.0 | 607 | 1 | LEPA_LACLA |
| 18 | 50 | 37.0 | 647 | 1 | SKO1_YEAST |
| 19 | 49 | 36.3 | 330 | 1 | GCAA_MOUSE |
| 20 | 49 | 36.3 | 336 | 1 | GCB_MOUSE |
| 21 | 49 | 36.3 | 399 | 1 | GCAH_MOUSE |
| 22 | 49 | 36.3 | 405 | 1 | GCBM_MOUSE |
| 23 | 48 | 35.6 | 437 | 1 | VIT1_CERCA |
| 24 | 48 | 35.6 | 777 | 1 | BAR1_HUMAN |
| 25 | 47 | 34.8 | 264 | 1 | H2D2_RAT |
| 26 | 47 | 34.8 | 327 | 1 | GC4_HUMAN |
| 27 | 47 | 34.8 | 600 | 1 | LEPA_RICPR |
| 28 | 47 | 34.8 | 621 | 1 | LEPA_THEMA |
| 29 | 47 | 34.8 | 3135 | 1 | S230_PLAFO |
| 30 | 46.5 | 34.4 | 454 | 1 | MUC_MESAU |
| 31 | 46 | 34.1 | 217 | 1 | NODB_RHIME |
| 32 | 46 | 34.1 | 284 | 1 | TRUA_STRCO |
| 33 | 46 | 34.1 | 319 | 1 | CDK4_XENLA |

RESULT 1

| ID | EPC_HUMAN | STANDARD; | PRT; | 428 AA. |
|----|------------------------------------------------------------------------|-----------|------|---------|
| DT | 21-JUL-1986 (Rel. 01, Created) | | | |
| DT | 21-JUL-1986 (Rel. 01, Last sequence update) | | | |
| DT | 20-AUG-2001 (Rel. 40, Last annotation update) | | | |
| DE | IG EPSILON CHAIN C REGION. | | | |
| GN | IGHE. | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=83168897; PubMed=6300763; | | | |
| RA | Seno M., Kurokawa T., Ono Y., Onda H., Sasada R., Igarashi K., | | | |
| RA | Kikuchi M., Sugino Y., Nishida Y., Honjo T.; | | | |
| RT | "Molecular cloning and nucleotide sequencing of human immunoglobulin | | | |
| RT | epsilon chain cDNA."; | | | |
| RL | Nucleic Acids Res. 11:719-726(1983). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=83001945; PubMed=6288268; | | | |
| RA | Max E.E., Battey J., Ney R., Kirsch I.R., Leder P.; | | | |
| RT | "Duplication and deletion in the human immunoglobulin epsilon genes."; | | | |
| RL | Cell 29:691-699(1982). | | | |
| RN | [3] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=84236029; PubMed=62341164; | | | |
| RA | Flanagan J.G., Rabbitts T.H.; | | | |
| RT | "The sequence of a human immunoglobulin epsilon heavy chain constant | | | |
| RT | region gene, and evidence for three non-allelic genes."; | | | |
| RL | EMBO J. 1:655-660(1982). | | | |
| RN | [4] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=84207910; PubMed=6327276; | | | |
| RA | Ueda S., Nakai S., Nishida Y., Hisajima H., Honjo T.; | | | |
| RT | "Long terminal repeat-like elements flank a human immunoglobulin | | | |
| RT | epsilon pseudogene that lacks introns."; | | | |
| RL | EMBO J. 1:1539-1544(1982). | | | |
| RN | [5] | | | |
| RP | PRELIMINARY SEQUENCE (MYELOMA PROTEIN ND). | | | |
| RA | Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.; | | | |
| RT | (In) Bach M.K. (eds.); | | | |
| RL | Immediate hypersensitivity: modern concepts and developments, pp.1-36, | | | |
| RL | Marcel Dekker, New York (1978). | | | |
| RN | [6] | | | |
| RP | SEQUENCE OF 1-40: 68-114 AND 427-428 FROM N.A. | | | |
| RX | MEDLINE=83065234; PubMed=6815656; | | | |
| RA | Kenten J.H., Moigaard H.V., Houghton M., Derbyshire R.B., Viney J., | | | |
| RA | Bell L.O., Gould H.J.; | | | |
| RT | "Cloning and sequence determination of the gene for the human | | | |
| RT | immunoglobulin epsilon chain expressed in a myeloma cell line."; | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982). | | | |
| RN | [7] | | | |

ALIGNMENTS

| 34 | 46 | 34.1 | 326 | 1 | GC2_HUMAN |
|----|------|------|------|---|------------|
| 35 | 46 | 34.1 | 329 | 1 | GCC_RAT |
| 36 | 46 | 34.1 | 602 | 1 | LEPA_CHLPN |
| 37 | 46 | 34.1 | 754 | 1 | ASPH_BOVIN |
| 38 | 46 | 34.1 | 1121 | 1 | YE06_YEAST |
| 39 | 45.5 | 33.7 | 765 | 1 | BAR1_MOUSE |
| 40 | 45.5 | 33.7 | 768 | 1 | BAR1_RAT |
| 41 | 45 | 33.3 | 63 | 1 | DN71_SULSO |
| 42 | 45 | 33.3 | 63 | 1 | DN72_SULSH |
| 43 | 45 | 33.3 | 63 | 1 | DN72_SULSO |
| 44 | 45 | 33.3 | 229 | 1 | NODB_RHILP |
| 45 | 45 | 33.3 | 290 | 1 | GC3_HUMAN |

P01859 homo sapien
P20762 rattus norv
Q92814 chlamydia p
Q28056 bos taurus
P32644 saccharomyc
O70445 mus musculu
Q92422 rattus norv
P80170 sulfolobus
O59632 sulfolobus
P39476 sulfolobus
P24150 rhizobium l
P01860 homo sapien

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 [1]
 RN SEQUENCE FROM N.A. (IMMUNOCYTOMA IR2).
 RP STRAIN=LOU/C/MSL;
 RX MEDLINE=83064537; PubMed=6292865;
 RA Hellman L., Pettersson U., Engstroem A., Karlsson T., Bennich H.;
 RT "Structure and evolution of the heavy chain from rat immunoglobulin
 E."; Nucleic Acids Res. 10:6041-6049(1982).
 RL [2]
 RN SEQUENCE OF 168-342 FROM N.A. (MYELOMA IR162).
 RP MEDLINE=83182019; PubMed=6820340;
 RX Kindsvogel W.R., Reddy E.P., Moore J.M., Faust C.H. Jr.;
 RA "A cloned cDNA probe for rat immunoglobulin epsilon heavy chain:
 RT construction, identification, and DNA sequence."; DNA 1:335-343(1982).
 RL [3]
 RN SEQUENCE OF 205-306 FROM N.A.
 RP MEDLINE=82174576; PubMed=6803238;
 RX Hellman L., Pettersson U., Bennich H.;
 RA "Characterization and molecular cloning of the mRNA for the heavy
 RT (epsilon) chain of rat immunoglobulin E."; Proc. Natl. Acad. Sci. U.S.A. 79:1264-1268(1982).
 RL [4]
 RN SEQUENCE OF 429 AA; 48671 MW; D2970B34EF8A72B0 CRC64;
 RP PIR: A02143; EHRT.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003597; Ig_CL.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig; 4.
 DR SMART: SM00407; IGcl; 1.
 DR SMART: SM00410; IG_Like; 3.
 DR PROSITE: PS00290; IG_MHC; 3.
 KW Immunoglobulin domain; Immunoglobulin C region.
 FT NON_TER 1
 FT CONFLICT 168 168 R -> N (IN REF. 2).
 FT CONFLICT 308 308 P -> L (IN REF. 2).
 SQ SEQUENCE 429 AA; 48671 MW; D2970B34EF8A72B0 CRC64;
 Query Match 48.9%; Score 66; DB 1; Length 429;
 Best Local Similarity 52.2%; Pred. No. 0.015;
 Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;
 [1]
 QY 2 GETYKSTVSHDPLPREVVRSTAK 24
 [1]
 DB 284 GEGYQCRVDHPHPKPIVRSTK 306
 [1]
 RESULT 4
 LEPA_HAEIN STANDARD; PRT; 598 AA.
 ID LEPA_HAEIN
 AC P43729;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE GTP-BINDING PROTEIN LEPA.
 GN LEPA OR HI0016.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OX NCBI_TaxID=727;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus
 RT influenzae Rd."; Science 269:496-512(1995).
 RL [1]
 RN SUBCELLULAR LOCATION: PRESENT IN THE CYTOPLASMIC MEMBRANE AND IS
 CC ALSO FOUND IN THE PERIPLASM (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC LEPA SUBFAMILY.
 CC [1]
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC [2]
 CC EMBL: U32687; AAC21694.1; -
 DR HSSP: P13551; ILEO.
 DR TIGR: HI0016;
 DR InterPro: IPR000795; GTP_EFTU.
 DR Pfam: PF00009; GTP_EFTU; 1.
 DR PROSITE: PS00301; EFATOR_GTP; 1.
 KW Membrane; GTP-binding; Complete proteome.
 FT NP_BIND 11 18 GTP (BY SIMILARITY).
 FT NP_BIND 77 81 GTP (BY SIMILARITY).
 FT NP_BIND 131 134 GTP (BY SIMILARITY).
 SQ SEQUENCE 598 AA; 66333 MW; FDDBIAC3516BF11E CRC64;
 Query Match 40.0%; Score 54; DB 1; Length 598;
 Best Local Similarity 46.4%; Pred. No. 1.4;
 Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;
 [1]
 QY 2 GETYKSTVSHDPLPREVVRSTAK 25
 [1]
 DB 68 GETYQLNFIDTPGHVDFSEVSRSLAAC 95
 [1]
 RESULT 5
 LEPA_PASMU STANDARD; PRT; 598 AA.
 ID LEPA_PASMU
 AC P57806;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE GTP-BINDING PROTEIN LEPA.
 GN LEPA OR PM0063.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OX NCBI_TaxID=747;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70."; Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC LEPA SUBFAMILY.
 CC [1]

DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IgC1; 2.
DR SMART; SM00410; Ig_Like; 2.
DR PROSITE; PS00290; Ig_MHC; 3.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
SQ SEQUENCE 450 AA; 4895 MW; 90460DA9D1012F5D CRC64;

Query Match 38.9%; Score 52.5; DB 1; Length 450;
Best Local Similarity 50.0%; Pred. No. 1.8;
Matches 12; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 2 GETYKSTVSHDPLPREVVRSIAK 24
||||| :|:|:| :|:|:
Db 405 GETYCVVAHESLPNRVTSVDK 428

RESULT 10
ALC2_HUMAN STANDARD; PRT; 340 AA.
ID ALC2_HUMAN
AC P01877;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE IG ALPHA-2 CHAIN C REGION.
DE IGH2.
GN IG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84130179; PubMed=6421489;
RA Flanagan J.G., Lefranc M.-P., Rabbitts T.H.;
RT "Mechanisms of divergence and convergence of the human immunoglobulin
alpha 1 and alpha 2 constant region gene sequences.";
RL Cell 36:681-688(1984).
RN [2]
RP SEQUENCE (BUT).
RX MEDLINE=78137069; PubMed=416441;
RA Torano A., Putnam F.W.;
RT "Complete amino acid sequence of the alpha 2 heavy chain of a human
IgA2 immunoglobulin of the A2m (2) allotype.";
RL Proc. Natl. Acad. Sci. U.S.A. 75:966-969(1978).
RN [3]
RP SEQUENCE (MYELOMA PROTEIN LAN).
RX MEDLINE=79180140; PubMed=286295;
RA Tazukida Y., Wang C.-C., Putnam F.W.;
RT "Structure of the A2m(1) allotype of human IgA -- a recombinant
molecule.";
RL Proc. Natl. Acad. Sci. U.S.A. 76:1104-1108(1979).
RN [4]
RP REVIEW.
RX MEDLINE=91054387; PubMed=2241915;
RA Kerr M.A.;
RT "The structure and function of human IgA.";
RL Biochem. J. 271:285-296(1990).
CC -1- FUNCTION: IG ALPHA IS THE MAJOR IMMUNOGLOBULIN CLASS IN BODY
SECRETIONS. IT MAY SERVE BOTH TO DEFEND AGAINST LOCAL INFECTION
AND TO PREVENT ACCESS OF FOREIGN ANTIGENS TO THE GENERAL
IMMUNOLOGIC SYSTEM.
CC -1- SUBUNIT: MONOMERIC OR POLYMERIC.
CC -1- MISCELLANEOUS: THE SEQUENCE OF THE A2M(1) ALLOTYPE IS SHOWN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; J00221; AAB59396.1; ALT_INIT.
DR PIR; A02172; A2HU.
DR PIR; B22360; B22360.
DR PIR; C22360; C22360.
DR MIW; 147000;
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00407; IgC1; 2.
DR SMART; SM00410; Ig_Like; 1.
DR PROSITE; PS00290; Ig_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DISULFID 26 85
FT DISULFID 101 101
FT FT
FT DISULFID 109 109
FT DISULFID 110 167
FT DISULFID 134 191
FT DISULFID 169 179
FT DISULFID 179 179
FT DISULFID 237 300
FT DISULFID 339 339
FT CARBOHYD 47 47
FT CARBOHYD 92 92
FT CARBOHYD 131 131
FT CARBOHYD 205 205
FT CARBOHYD 327 327
FT VARIANT 93 93
FT FT
FT VARIANT 102 102
FT FT
FT VARIANT 279 279
FT FT
FT VARIANT 296 296
FT FT
FT VARIANT 326 326
FT FT
FT VARIANT 335 335
FT FT
SQ SEQUENCE 340 AA; 36508 MW; 98922700756F3276 CRC64;

Query Match 37.8%; Score 51; DB 1; Length 340;
Best Local Similarity 39.1%; Pred. No. 2.2;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 GETYKSTVSHDPLPREVVRSIAK 24
||||| :|:|:| :|:|:
Db 186 GETFTCTAAHPKLPTANITK 208

RESULT 11
MUCB_HUMAN STANDARD; PRT; 391 AA.
ID MUCB_HUMAN
AC P04220;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE IG MU HEAVY CHAIN DISEASE PROTEIN (BOT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=84184186; PubMed=6425189;
RA Barnikol-Watanabe S., Mihaesco E., Mihaesco C., Barnikol H.U.,
RA Hilschmann N.;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 13:08:13 ; Search time 79.15 Seconds
(without alignments)
46.201 Million cell updates/sec

Title: US-09-701-623c-84
Perfect score: 135
Sequence: 1 CGETYKSTVSHPDLPREVVRSTAKC 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-----------|--------------------|
| 1 | 54 | 40.0 | 137 | 2 Q57464 | Q57464 photobacter |
| 2 | 53 | 39.3 | 684 | 13 Q90544 | Q90544 ginglymosto |
| 3 | 52 | 38.5 | 394 | 5 O17620 | O17620 caenorhabdi |
| 4 | 52 | 38.5 | 648 | 2 Q9A94 | Q9a9f4 caulobacter |
| 5 | 51 | 37.8 | 93 | 7 Q31260 | Q31260 rattus norv |
| 6 | 51 | 37.8 | 341 | 11 Q08284 | Q08284 mus musculu |
| 7 | 51 | 37.8 | 369 | 2 Q9CM48 | Q9cm48 pasteurella |
| 8 | 51 | 37.8 | 375 | 4 Q9BS21 | Q9bsz1 homo sapien |
| 9 | 51 | 37.8 | 416 | 4 Q9NPP6 | Q9npp6 homo sapien |
| 10 | 51 | 37.8 | 557 | 11 Q9K47 | Q9k47 mus musculu |
| 11 | 51 | 37.8 | 597 | 2 Q9KPE0 | Q9kpb0 vibrio chol |
| 12 | 51 | 37.8 | 597 | 4 Q9BU10 | Q9bul0 homo sapien |
| 13 | 51 | 37.8 | 597 | 4 Q9BQ88 | Q9bqb8 homo sapien |
| 14 | 51 | 37.8 | 882 | 5 Q9NE77 | Q9ne67 leishmania |
| 15 | 50.5 | 37.4 | 893 | 5 O02099 | O02099 caenorhabdi |
| 16 | 50 | 37.0 | 245 | 2 Q9EV20 | Q9ev20 listeria mo |
| 17 | 50 | 37.0 | 599 | 2 Q9I568 | Q9i568 pseudomonas |
| 18 | 50 | 37.0 | 610 | 2 Q99ZV8 | Q99zv8 streptococc |
| 19 | 49 | 36.3 | 520 | 10 O48786 | O48786 arabidopsis |

| | | | | | |
|----|------|------|------|-----------|--------------------|
| 20 | 49 | 36.3 | 1105 | 5 Q9VM44 | Q9vn44 drosophila |
| 21 | 48.5 | 35.9 | 279 | 11 Q9D9F8 | Q9d9f8 mus musculu |
| 22 | 48 | 35.6 | 335 | 11 Q08283 | Q08283 mesocricetu |
| 23 | 48 | 35.6 | 425 | 11 Q9DB22 | Q9db22 mus musculu |
| 24 | 48 | 35.6 | 962 | 5 Q9GYU5 | Q9gyu5 trypanosoma |
| 25 | 48 | 35.6 | 1048 | 13 P79954 | P79954 xenopus lae |
| 26 | 47 | 34.8 | 217 | 7 Q30827 | Q30827 ovis aries |
| 27 | 47 | 34.8 | 245 | 7 Q31270 | Q31270 rattus norv |
| 28 | 47 | 34.8 | 261 | 7 Q9TQA7 | Q9tqa7 rattus norv |
| 29 | 47 | 34.8 | 264 | 7 Q9TQA5 | Q9tqa5 rattus norv |
| 30 | 47 | 34.8 | 387 | 5 O17621 | O17621 caenorhabdi |
| 31 | 47 | 34.8 | 479 | 10 Q9LWH4 | Q9lwh4 oryza sativ |
| 32 | 47 | 34.8 | 607 | 2 Q99TR4 | Q99tr4 staphylococ |
| 33 | 47 | 34.8 | 3134 | 5 Q25994 | Q25994 plasmodium |
| 34 | 46 | 34.1 | 176 | 2 Q9R8Y8 | Q9r8y8 rhizobium s |
| 35 | 46 | 34.1 | 176 | 2 Q9R8Y2 | Q9r8y2 rhizobium s |
| 36 | 46 | 34.1 | 177 | 2 Q9R8X6 | Q9r8x6 rhizobium s |
| 37 | 46 | 34.1 | 178 | 2 Q923P1 | Q923p1 rhizobium s |
| 38 | 46 | 34.1 | 178 | 2 Q9R8Z0 | Q9r8z0 rhizobium s |
| 39 | 46 | 34.1 | 178 | 2 Q9R8Y9 | Q9r8y9 rhizobium s |
| 40 | 46 | 34.1 | 178 | 2 Q9R8Y5 | Q9r8y5 rhizobium s |
| 41 | 46 | 34.1 | 178 | 2 Q9R8Y4 | Q9r8y4 rhizobium s |
| 42 | 46 | 34.1 | 178 | 2 Q9R8Y0 | Q9r8y0 rhizobium s |
| 43 | 46 | 34.1 | 178 | 2 Q9R8X8 | Q9r8x8 rhizobium s |
| 44 | 46 | 34.1 | 178 | 2 Q9R8X4 | Q9r8x4 rhizobium s |
| 45 | 46 | 34.1 | 180 | 2 Q923N8 | Q923n8 rhizobium s |

ALIGNMENTS

RESULT 1

Q57464 ID Q57464 PRELIMINARY; PRT: 137 AA.
AC Q57464;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE RPOE, LEPA AND ORF 2, 3, 4 GENES (FRAGMENT).
GN LEPA.
OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae;
OC Photobacterium.
OX NCBI_TaxID=74109;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SS9;
RA Chi E., Bartlett D.H.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SS9;
RA Chi E., Bartlett D.H.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO GTP-BINDING ELONGATION FACTOR FAMILY.
DR EMBL; L41688; AAA65230.1; -;
DR EMBL; L41667; AAB31927.1; -;
DR HSSP; P13551; IFNM.
DR InterPro; IPR000795; GTP_EFTU.
DR Pfam; PF00009; GTP_EFTU; 1.
DR PROSITE; PS00301; EFACOR_GTP; 1.
FT GTP-binding; Protein biosynthesis.
FT NON_TER 137 137
SQ SEQUENCE 137 AA; 15016 MW; 7FC5E05FD6A1F3E CRC64;

Query Match 40.0%; Score 54; DB 2; Length 137;
Best Local Similarity 46.4%; Pred. No. 2.2;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

Qy 2 GETYK----STVSHPDLPREVVRSTAKC 25

||||: | | | | | | | | | |

Db 68 GETYQLNFIDFGHDFSYEVSRLAAC 95

```

RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA Smaidon N., Smith A., Sonhammer E., Staden K., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
DR EMBL; Z92827; CAB07326.1; -.
SQ SEQUENCE 394 AA; 44297 MW; D442A972AE24451 CRC64;

Query Match 38.5%; Score 52; DB 5; Length 394;
Best Local Similarity 52.9%; Pred. No. 13;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 9 VSHPDLPREVVRISIAK 25
Db 60 LKHPNLPKNVVIKASC 76

RESULT 4
Q9A9F4 PRELIMINARY; PRT; 648 AA.
ID O9A9F4
AC O9A9F4:
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE GTP-BINDING PROTEIN LEPA.
GN CC1034.
OS Caulobacter crescentus
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RX MEDLINE=2117369; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005781; AAK23018.1; -.
DR TIGR; CC1034; -.
KW Complete proteome.
SQ SEQUENCE 648 AA; 71538 MW; 3B62E3E2270A9249 CRC64;

Query Match 38.5%; Score 52; DB 2; Length 648;
Best Local Similarity 46.4%; Pred. No. 22;
Matches 13; Conservative 1; Mismatches 10; Indels 4; Gaps 1;

Qy 2 GETY-----KSTVSHPDLPREVVRISIAK 25
Db 119 GETYILNLMDDTPGHVDFAYEVSRLAAC 146

RESULT 5
Q31260 PRELIMINARY; PRT; 93 AA.
ID Q31260
AC Q31260:
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE RAT MHC CLASS II RT1.B-BETA2 (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=KIDNEY;
RA MEDLINE=86084127; PubMed=3865893;
RX Scholler J., Lernmark A.;
RT "Isolation of a rat immune response gene identical to an alleged mouse
a class II beta-chain pseudogene.";
RL Immunogenetics 22:601-608(1985).
CC -!- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
DOMAIN.
DR EMBL: M15561; AAA68203.1; -;
DR HSP; P06343; IIAK.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00407; Ig_c1; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
KW MHC.
FT NON_TER 1
FT NON_TER 93 93
SQ SEQUENCE 93 AA; 10605 MW; 308922D17EBD9F38 CRC64;

Query Match 37.8%; Score 51; DB 7; Length 93;
Best Local Similarity 52.9%; Pred. No. 4;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2 GETYKSTVSHPDLPREV 18
I : I I I I I I I
DB 73 GDVYSLVEHPSLPREV 89

RESULT 6
Q08284 PRELIMINARY; PRT; 341 AA.
AC Q08284;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE FIBRINOGEN ALPHA CHAIN (FRAGMENT).
GN FGA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=LIVER;
RA Murakawa M.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=LIVER;
RX MEDLINE=93269219; PubMed=8497848;
RA Murakawa M., Okamura T., Kamura T., Shibuya T., Harada M., Niho Y.;
RT "Diversity of primary structures of the carboxy-terminal regions of
mammalian fibrinogen A alpha-chains. Characterization of the partial
nucleotide and deduced amino acid sequences in five mammalian species;
rhesus monkey, pig, dog, mouse and Syrian hamster.";
Thromb. Haemost. 69:351-360(1993).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
(ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR EMBL: D43759; BAA07816.1; -;
DR HSP; P02671; IFZA.

DR MGD; MGI:1316726; Fga.
KW Blood coagulation; Plasma; Platelet; Repeat.
RN NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=KIDNEY;
RA MEDLINE=86084127; PubMed=3865893;
RX Scholler J., Lernmark A.;
RT "Isolation of a rat immune response gene identical to an alleged mouse
a class II beta-chain pseudogene.";
RL Immunogenetics 22:601-608(1985).
CC -!- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
DOMAIN.
DR EMBL: M15561; AAA68203.1; -;
DR HSP; P06343; IIAK.
DR InterPro: IPR003597; Ig_c1.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00407; Ig_c1; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
KW MHC.
FT NON_TER 1
FT NON_TER 93 93
SQ SEQUENCE 93 AA; 10605 MW; 308922D17EBD9F38 CRC64;

Query Match 37.8%; Score 51; DB 11; Length 341;
Best Local Similarity 45.5%; Pred. No. 16;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 CGETYKSTVSHPDLPREVRSI 22
I : I I I I I I I
DB 229 CSKTTTKVTGPDGRREVVKREV 250

RESULT 7
Q9CM48 PRELIMINARY; PRT; 369 AA.
AC Q9CM48;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE METB.
GN METB OR PM0995.
OC Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RL "Complete genomic sequence of Pasteurella multocida Pm70.";
Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL: AE006138; AAK03079.1; -;
DR InterPro: IPR000277; Cys_Met_Meta_PP.
DR Pfam: PF01053; Cys_Met_Meta_PP; 1.
DR PROSITE: PS00868; CYS_MET_METAB_PP; 1.
KW Complete proteome.
SQ SEQUENCE 369 AA; 40565 MW; EB8654C9FB69C5F0 CRC64;

Query Match 37.8%; Score 51; DB 2; Length 369;
Best Local Similarity 37.5%; Pred. No. 17;
Matches 9; Conservative 5; Mismatches 8; Indels 2; Gaps 1;

QY 4 TYKSTVSHPDLPREVRSIAK 25
I : I I I I I I I
DB 320 TYPATQTHMDIPEERARGVCNC 343

RESULT 8
Q9BSZ1 PRELIMINARY; PRT; 375 AA.
AC Q9BSZ1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

```
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:10455).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPHOMA;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC004476; AA04476.1; -.
SQ SEQUENCE 375 AA; 41314 MW; B1A0A0998F473619 CRC64;

Query Match 37.8%; Score 51; DB 4; Length 375;
Best Local Similarity 39.1%; Pred. No. 17;
Matches 9; Conservative 7; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GETYKSTVSHPDLPREVRSIAK 24
Db 220 GERFTCTVHTDLPSPKQTISR 242
|||||
220 GERFTCTVHTDLPSPKQTISR 242

RESULT 9
ID Q9NPP6 PRELIMINARY; PRT; 416 AA.
AC Q9NPP6;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE IMMUNOGLOBULIN HEAVY CHAIN VARIANT (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pluvinet R., Estivill X., Escarceller M., Sumoy L.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Auffray C., Ansorge W., Ballabio A., Estivill X., Gibson K.,
RA Lehrach H., Poustka A., Lundeberg J.;
RT "The European IMAGE Consortium for integrated Molecular analysis of
RT human gene transcripts."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
CC DOMAIN.
DR EMBL: AL389978; CAB97534.1; -.
DR InterPro: IPR003600; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 4.
DR SMART: SM00406; IGV; 1.
DR SMART: SM00410; IG_Like; 3.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_2.
FT NON_TER 1
SQ SEQUENCE 416 AA; 44786 MW; 8C41708BB8B4587 CRC64;

Query Match 37.8%; Score 51; DB 4; Length 416;
Best Local Similarity 39.1%; Pred. No. 19;
Matches 9; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GETYKSTVSHPDLPREVRSIAK 24
Db 262 GETFTCTAAHPKLTPLTANITK 284
|||||
262 GETFTCTAAHPKLTPLTANITK 284

RESULT 10
Q99K47 .
ID Q99K47 PRELIMINARY; PRT; 557 AA.

Query Match 37.8%; Score 51; DB 2; Length 597;
Best Local Similarity 42.9%; Pred. No. 28;
Matches 12; Conservative 3; Mismatches 9; Indels 4; Gaps 1;

Qy 2 GETYKSTVSHPDLPREVRSIAK 25
```

```
AC Q99K47;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE FIBRINOGEN A ALPHA POLYPEPTIDE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MAMMARY TUMOR;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC005467; AA05467.1; -.
SQ SEQUENCE 557 AA; 61325 MW; C47F496D1BA432DE CRC64;

Query Match 37.8%; Score 51; DB 11; Length 557;
Best Local Similarity 45.5%; Pred. No. 26;
Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHPDLPREVRSI 22
Db 408 CSKTTTKVTGPDGRREVVEV 429
|||||
408 CSKTTTKVTGPDGRREVVEV 429

RESULT 11
ID Q9KPB0 PRELIMINARY; PRT; 597 AA.
AC Q9KPB0;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE GTP-BINDING PROTEIN LEPA.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -1- SIMILARITY: TO GTP-BINDING ELONGATION FACTOR FAMILY.
DR EMBL: AE004316; AAF95605.1; -.
DR TIGR: VC2463; -.
DR InterPro: IPR002106; AA_trna_ligase_II.
DR InterPro: IPR000795; GTP_EFTU.
DR InterPro: IPR001806; Ras_trnsfrmg.
DR Pfam: PF00009; GTP_EFTU; 1.
DR PRINTS: PR00315; ELONGATNFCT.
DR PRINTS: PR00449; RASTRNSFRMG.
DR PROSITE: PS00339; AA_TRNA_LIGASE_II_2; UNKNOWN_1.
DR PROSITE: PS00301; EFATOR_GTP; 1.
SQ Complete proteome; GTP-binding.
KW SEQUENCE 597 AA; 66029 MW; 754C1EE2E3DF0DDE CRC64;

Query Match 37.8%; Score 51; DB 2; Length 597;
Best Local Similarity 42.9%; Pred. No. 28;
Matches 12; Conservative 3; Mismatches 9; Indels 4; Gaps 1;

Qy 2 GETYKSTVSHPDLPREVRSIAK 25
```

[illegible]

| Case | Age (years) | Sex | Site | Pathologic findings | Outcome |
|------|-------------|-----|--------|---------------------|------------------|
| 1 | 65 | M | Rectum | Adenocarcinoma | Survived 5 years |
| 2 | 72 | F | Rectum | Adenocarcinoma | Survived 3 years |
| 3 | 68 | M | Rectum | Adenocarcinoma | Survived 4 years |
| 4 | 70 | F | Rectum | Adenocarcinoma | Survived 2 years |
| 5 | 63 | M | Rectum | Adenocarcinoma | Survived 6 years |
| 6 | 75 | F | Rectum | Adenocarcinoma | Survived 1 year |
| 7 | 60 | M | Rectum | Adenocarcinoma | Survived 3 years |
| 8 | 71 | F | Rectum | Adenocarcinoma | Survived 2 years |
| 9 | 66 | M | Rectum | Adenocarcinoma | Survived 4 years |
| 10 | 73 | F | Rectum | Adenocarcinoma | Survived 1 year |
| 11 | 69 | M | Rectum | Adenocarcinoma | Survived 3 years |
| 12 | 74 | F | Rectum | Adenocarcinoma | Survived 2 years |
| 13 | 67 | M | Rectum | Adenocarcinoma | Survived 4 years |
| 14 | 76 | F | Rectum | Adenocarcinoma | Survived 1 year |
| 15 | 64 | M | Rectum | Adenocarcinoma | Survived 3 years |
| 16 | 77 | F | Rectum | Adenocarcinoma | Survived 2 years |
| 17 | 62 | M | Rectum | Adenocarcinoma | Survived 4 years |
| 18 | 78 | F | Rectum | Adenocarcinoma | Survived 1 year |
| 19 | 61 | M | Rectum | Adenocarcinoma | Survived 3 years |
| 20 | 79 | F | Rectum | Adenocarcinoma | Survived 2 years |

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 4, 2002, 12:56:35 ; Search time 38.04 Seconds
(without alignments)
14.789 Million cell updates/sec

Title: us-09-701-623c-84

Perfect score: 135

Sequence: 1 CGETKSTVSHPDLPREVVRSTAKC 25

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents_AA:*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*

2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*

3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*

4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*

5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*

6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 94 | 69.6 | 25 | 3 | US-09-100-414B-95 |
| 2 | 94 | 69.6 | 25 | 4 | US-09-303-323-95 |
| 3 | 94 | 69.6 | 42 | 3 | US-09-100-414B-98 |
| 4 | 94 | 69.6 | 42 | 3 | US-09-100-414B-99 |
| 5 | 94 | 69.6 | 42 | 3 | US-09-100-414B-100 |
| 6 | 94 | 69.6 | 42 | 4 | US-09-303-323-98 |
| 7 | 94 | 69.6 | 42 | 4 | US-09-303-323-99 |
| 8 | 94 | 69.6 | 42 | 4 | US-09-303-323-100 |
| 9 | 94 | 69.6 | 45 | 3 | US-09-100-414B-101 |
| 10 | 94 | 69.6 | 45 | 4 | US-09-303-323-101 |
| 11 | 94 | 69.6 | 46 | 3 | US-09-100-414B-96 |
| 12 | 94 | 69.6 | 46 | 3 | US-09-100-414B-97 |
| 13 | 94 | 69.6 | 45 | 4 | US-09-303-323-96 |
| 14 | 94 | 69.6 | 45 | 4 | US-09-303-323-97 |
| 15 | 94 | 69.6 | 63 | 3 | US-09-100-414B-102 |
| 16 | 94 | 69.6 | 63 | 4 | US-09-303-323-102 |
| 17 | 82 | 60.7 | 426 | 1 | US-08-336-583-2 |
| 18 | 82 | 60.7 | 426 | 5 | PCT-US95-13795-2 |
| 19 | 71 | 52.6 | 106 | 2 | US-08-232-539D-54 |
| 20 | 71 | 52.6 | 113 | 2 | US-08-232-539D-56 |
| 21 | 60.5 | 44.8 | 119 | 2 | US-08-464-025A-1 |
| 22 | 58 | 43.0 | 22 | 2 | US-08-232-539D-19 |
| 23 | 58 | 43.0 | 24 | 2 | US-08-232-539D-20 |
| 24 | 58 | 43.0 | 56 | 2 | US-08-232-539D-18 |
| 25 | 54 | 40.0 | 561 | 3 | US-09-192-545-2 |
| 26 | 51 | 37.8 | 96 | 6 | 5284931-11 |
| 27 | 51 | 37.8 | 106 | 1 | US-08-399-106A-7 |

| | | | | | | |
|----|----|------|-----|---|------------------|--------------------|
| 28 | 51 | 37.8 | 106 | 1 | US-08-433-105A-7 | Sequence 7, Appli |
| 29 | 51 | 37.8 | 106 | 2 | US-08-434-869A-7 | Sequence 7, Appli |
| 30 | 51 | 37.8 | 483 | 4 | US-09-049-672A-5 | Sequence 5, Appli |
| 31 | 51 | 37.8 | 499 | 4 | US-09-049-672A-1 | Sequence 1, Appli |
| 32 | 51 | 37.8 | 504 | 1 | US-07-932-915-2 | Sequence 2, Appli |
| 33 | 51 | 37.8 | 504 | 5 | PCT-US91-05826-2 | Sequence 2, Appli |
| 34 | 49 | 36.3 | 357 | 3 | US-08-630-172-20 | Sequence 20, Appli |
| 35 | 49 | 36.3 | 357 | 4 | US-09-375-419-20 | Sequence 20, Appli |
| 36 | 49 | 36.3 | 367 | 3 | US-08-630-172-19 | Sequence 19, Appli |
| 37 | 49 | 36.3 | 367 | 4 | US-09-375-419-19 | Sequence 19, Appli |
| 38 | 49 | 36.3 | 378 | 3 | US-08-630-172-21 | Sequence 21, Appli |
| 39 | 49 | 36.3 | 378 | 4 | US-09-375-419-21 | Sequence 21, Appli |
| 40 | 49 | 36.3 | 386 | 4 | US-08-875-533-42 | Sequence 42, Appli |
| 41 | 49 | 36.3 | 392 | 4 | US-08-875-533-41 | Sequence 41, Appli |
| 42 | 49 | 36.3 | 396 | 4 | US-08-875-533-68 | Sequence 68, Appli |
| 43 | 49 | 36.3 | 410 | 3 | US-08-630-172-17 | Sequence 17, Appli |
| 44 | 49 | 36.3 | 410 | 4 | US-09-375-419-17 | Sequence 17, Appli |
| 45 | 49 | 36.3 | 425 | 5 | PCT-US95-03866-8 | Sequence 8, Appli |

ALIGNMENTS

RESULT 1
US-09-100-414B-95
: Sequence 95, Application US/09100414B
: Patent No. 6025468
: GENERAL INFORMATION:
: APPLICANT: Wang, Chang Yi
: TITLE OF INVENTION: NOVEL LHRH PEPTIDE
: TITLE OF INVENTION: IMMUNOGENS
: NUMBER OF SEQUENCES: 106
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Morgan & Finnegan, L.L.P.
: STREET: 345 Park Avenue
: CITY: New York
: STATE: NY
: COUNTRY: USA
: ZIP: 10154-0054
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC Windows
: SOFTWARE: Word 97
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/100,414B
: FILING DATE: 20-JUNE-1998
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Maria H. Lin
: REGISTRATION NUMBER: 29,323
: REFERENCE/DOCKET NUMBER: 1151-4157
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-758-4800
: TELEFAX: 212-751-6849
: INFORMATION FOR SEQ ID NO: 95:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 25 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-09-100-414B-95

Query Match 69.6%; Score 94; DB 3; Length 25;
Best Local Similarity 64.0%; Pred. No. 7.4e-09;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 CGETKSTVSHPDLPREVVRSTAKC 25

DB 1 CGETQSRVTHPLPALMRSTTKC 25

RESULT 2
US-09-303-323-95
; Sequence 95, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang YI
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 95:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-95

Query Match 69.6%; Score 94; DB 4; Length 25;
Best Local Similarity 64.0%; Pred. No. 7.4e-09;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHDLPREVVRSIAC 25
||||| :||| :||| :||| :|||
Db 1 CGETYQSRVTHPLPRALMRSTTK 25

RESULT 3
US-09-100-414B-98
; Sequence 98, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang YI
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows

; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-98

Query Match 69.6%; Score 94; DB 3; Length 42;
Best Local Similarity 64.0%; Pred. No. 1.4e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHDLPREVVRSIAC 25
||||| :||| :||| :||| :|||
Db 18 CGETYQSRVTHPLPRALMRSTTK 42

RESULT 4
US-09-100-414B-99
; Sequence 99, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang YI
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-99

Query Match 69.6%; Score 94; DB 3; Length 42;
Best Local Similarity 64.0%; Pred. No. 1.4e-08;

Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHPDLPREVVRSTKAC 25
|||||:|:|:|:|:|:|:|:|:|
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 5

US-09-100-414B-100
; Sequence 100, Application US/09100414B
; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100.414B
; FILING DATE: 20-JUNE-1998
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-100-414B-100

Query Match 69.6%; Score 94; DB 3; Length 42;
Best Local Similarity 64.0%; Pred. No. 1.4e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHPDLPREVVRSTKAC 25
|||||:|:|:|:|:|:|:|:|:|
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 6

US-09-303-323-98
; Sequence 98, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303.323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100.414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 98:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-98

Query Match 69.6%; Score 94; DB 4; Length 42;
Best Local Similarity 64.0%; Pred. No. 1.4e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHPDLPREVVRSTKAC 25
|||||:|:|:|:|:|:|:|:|:|
Db 18 CGETYQSRVTHPLPALMRSTTKC 42

RESULT 7

US-09-303-323-99
; Sequence 99, Application US/09303323
; Patent No. 6228987
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303.323
; FILING DATE: 30-APR-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100.414
; FILING DATE: 20-JUNE-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 99:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-99

Query Match 69.6%; Score 94; DB 4; Length 42;
Best Local Similarity 64.0%; Pred. No. 1.4e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHDPLPREVVRSIAC 25
Db 18 CGETYQSRVTHPLPRALMRSTTK 42

RESULT 8

US-09-303-323-100
; Sequence 100, Application US/09303323
; Patent No. 6228987

; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/303,323
; FILING DATE: 30-APR-1999

CLASSIFICATION:

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/100,414
; FILING DATE: 20-JUNE-1998

; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800

; TELEFAX: 212-751-6849

; INFORMATION FOR SEQ ID NO: 100:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 42 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-303-323-100

Query Match 69.6%; Score 94; DB 4; Length 42;
Best Local Similarity 64.0%; Pred. No. 1.4e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHDPLPREVVRSIAC 25
Db 18 CGETYQSRVTHPLPRALMRSTTK 42

RESULT 9

US-09-100-414B-101
; Sequence 101, Application US/09100414B

; Patent No. 6025468
; GENERAL INFORMATION:
; APPLICANT: Wang, Chang Yi
; TITLE OF INVENTION: NOVEL LHRH PEPTIDE
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10154-0054

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC Windows
; SOFTWARE: Word 97

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/100,414B
; FILING DATE: 20-JUNE-1998

CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:
; NAME: Maria H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4157

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-758-4800

; TELEFAX: 212-751-6849

; INFORMATION FOR SEQ ID NO: 101:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 45 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-09-100-414B-101

Query Match 69.6%; Score 94; DB 3; Length 45;
Best Local Similarity 64.0%; Pred. No. 1.5e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHDPLPREVVRSIAC 25

Db 21 CGETYQSRVTHPLPRALMRSTTK 45

RESULT 10

US-09-303-323-101

; Sequence 101, Application US/09303323

; Patent No. 6228987

; GENERAL INFORMATION:

; APPLICANT: Wang, Chang Yi

; TITLE OF INVENTION: NOVEL LHRH PEPTIDE

; NUMBER OF SEQUENCES: 106

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Morgan & Finnegan, L.L.P.

; STREET: 345 Park Avenue

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10154-0054

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC Windows

; SOFTWARE: Word 97

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/303,323

; FILING DATE: 30-APR-1999

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 96:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-96

Query Match 69.6%; Score 94; DB 4; Length 46;
Best Local Similarity 64.0%; Pred. No. 1.6e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHDPLPREVRSIAKC 25
|||||:|:|:|:|:|:|:|:|:|
Db 22 CGETYQSRVTHPLPRALMRSTTKC 46

RESULT 14
US-09-303-323-97
Sequence 97, Application US/09303323
Patent No. 6228987
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/303,323
FILING DATE: 30-APR-1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/100,414
FILING DATE: 20-JUNE-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 46 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-303-323-97

Query Match 69.6%; Score 94; DB 4; Length 46;
Best Local Similarity 64.0%; Pred. No. 1.6e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHDPLPREVRSIAKC 25
|||||:|:|:|:|:|:|:|:|:|
Db 22 CGETYQSRVTHPLPRALMRSTTKC 46

RESULT 15
US-09-100-414B-102
Sequence 102, Application US/09100414B
Patent No. 6025468
GENERAL INFORMATION:
APPLICANT: Wang, Chang Yi
TITLE OF INVENTION: NOVEL LHRH PEPTIDE
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morgan & Finnegan, L.L.P.
STREET: 345 Park Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10154-0054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC Windows
SOFTWARE: Word 97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/100,414B
FILING DATE: 20-JUNE-1998
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maria H. Lin
REGISTRATION NUMBER: 29,323
REFERENCE/DOCKET NUMBER: 1151-4157
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-758-4800
TELEFAX: 212-751-6849
INFORMATION FOR SEQ ID NO: 102:
SEQUENCE CHARACTERISTICS:
LENGTH: 63 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-100-414B-102

Query Match 69.6%; Score 94; DB 3; Length 63;
Best Local Similarity 64.0%; Pred. No. 2.3e-08;
Matches 16; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CGETYKSTVSHDPLPREVRSIAKC 25
|||||:|:|:|:|:|:|:|:|:|
Db 39 CGETYQSRVTHPLPRALMRSTTKC 63

Search completed: March 4, 2002, 12:56:35
Job time: 257 sec

THIS PAGE BLANK (USPTO)